1-1-2012

Exploring the Role of Wolbachia Endobacteria in the Biology of Filarial Nematode Parasites

Samantha McNulty
Washington University in St. Louis

WASHINGTON UNIVERSITY IN SAINT LOUIS

Division of Biology and Biomedical Sciences

Molecular Microbiology and Microbial Pathogenesis

Dissertation Examination Committee:
Gary J. Weil, Co-Chair
Peter U. Fischer, Co-Chair
Douglas E. Berg
Stephen M. Beverley
Daniel E. Goldberg
Jeffrey I. Gordon
Makedonka Mitreva

Exploring the Role of *Wolbachia* Endobacteria in the Biology of Filarial Nematode Parasites

by

Samantha Nichole McNulty

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2012
Saint Louis, Missouri

ABSTRACT

Filarial nematodes are vector borne parasitic worms that cause a variety of disfiguring and disabling diseases, including lymphatic filariasis and onchocerciasis. Many filarial species require *Wolbachia* endobacteria (family: Rickettsiaceae) to carry out their life cycle. Studies using antibiotics to target the endobacteria, thereby interfering with worm fertility and viability, have generated interest in using *Wolbachia* as an antifilarial drug target. However, the exact mechanisms underpinning this interesting mutualistic interaction are poorly understood.

Wolbachia-dependence is not ubiquitous in the filarial family. Some species are able to survive in the absence of an endosymbiont. The inconsistent patterns of *Wolbachia*-dependence and independence seen in filarial nematodes may be explained by two hypotheses. Following infection with the endobacteria, reductive evolution could have removed redundant genes or pathways present in both partners. Thus, deletions in the worm's genome would render it dependent on *Wolbachia* for vital gene products. Conversely, *Wolbachia*-dependent species could re-acquire vital genes from the endosymbiont by horizontal gene transfer, rendering the bacteria expendable.

Mitochondria and *Wolbachia* are co-transmitted vertically from mother to offspring, therefore the mitochondrial genome (mtDNA) is particularly sensitive to evolutionary pressures exerted by the endosymbiont. *Wolbachia* is also thought to be closely related to the mitochondrial progenitor, so they may overlap in function (e.g., energy production). In order to address our first hypothesis, we sequenced the mitochondrial genomes of several species of *Wolbachia*-dependent and independent filarial nematodes in hopes of finding some degeneracy in the mtDNA of the *Wolbachia*-

dependent species. Our studies have shown that the mtDNA of all examined species encodes the same 12 protein coding genes, 2 ribosomal RNA genes and 22 transfer RNA genes. Despite a careful analysis, no sequence-level differences were observed between the mtDNA of infected and uninfected species.

In order to address our second hypothesis, we surveyed the genomes of two *Wolbachia*-independent filarial species, *Acanthocheilonema viteae* and *Onchocerca flexuosa*, in search of evidence of horizontal gene transfer from *Wolbachia*. Many genomic fragments containing regions with high homology to *Wolbachia* sequences were identified. Follow-up transcriptomic and proteomic analyses in *O. flexuosa* have shown that *Wolbachia*-like sequnces are expressed at the RNA and protein levels. Imaging studies indicate that *Wolbachia*-like RNAs are mainly produced in tissues known to harbor *Wolbachia* in infected species, while a *Wolbachia*-like protein was found nearby but not in the same tissue.

This project has produced a vast amount of data that will be useful to the filariasis research community, including the mtDNA sequences of five filarial species, genomic sequences from *A. viteae* and *O. flexuosa*, transcriptomic sequences from *O. flexuosa* and a survey of the *O. flexuosa* adult worm proteome. Our results have verified a longstanding hypothesis that the ancestor(s) of many *Wolbachia*-free filarial nematode species was colonized in the distant past despite the present lack of endobacteria. Future studies may prove that horizontally transferred bacterial genes are necessary for the survival of *Wolbachia*-free filarial worms that would otherwise require *Wolbachia* for reproduction and development.

ACKNOWLEDGEMENTS

"No duty is more urgent than that of returning thanks." - Anonymous

I lack the space on these few pages to adequately express my gratitude to those who have been instrumental in the completion of this work. First and foremost, I would like to thank my incredible mentors, Gary Weil and Peter Fischer. Their complementary medical and biological perspectives have been invaluable in my education as a scientist. Thank you both for putting up with my fumblings as a beginner and granting me room for independence and creativity as I grew in experience.

I thank all members of the Weil Lab, past and present, for their role in my training. I've learned so much from each and every one of you, both scientifically and personally. Thank you for your help and for your friendship.

My thesis committee has played a tremendously important role in shaping my projects, not only through their insightful comments but through their willingness to collaborate and share resources. This amazing group of scientists ensured that I never wanted for anything, whether it be a reagent, borrowed equipment or simply a fresh perspective. Between the provisions of my mentors and committee, my projects were never limited in scope.

Great science is never done in isolation. I have been fortunate to collaborate with groups at The Genome Institute at Washington University, New England Biolabs (Ipswich, MA), The Bernard Nocht Institute for Tropical Medicine (Hamburg, Germany),

and the University of North Dakota. I thank these fine scientists for sharing their resources and their expertise.

I am grateful to have been awarded the graduate student training grant in cellular and molecular biology (NIH T32 GM07067, 9/1/05 to 9/31/08) and the infectious disease/mechanisms microbial pathogenesis training grant (T32-AI007172, 9/1/08 to 8/31/10).

Finally, I thank my family. Any success I have achieved in life is due, in large part, to the love and support of my parents, Cletus and Johanna Piper and Sharon and Bill Steele. My husband, Nate McNulty, is my best friend, my classmate, my teacher, my technician, my maintenance programmer, my everything. I would have given up one hundred times without his encouragement.

I spent the last year battling to complete and defend my thesis prior to the birth of my first child. On October 14, 2011 at 3:37am, I lost this battle to a small but formidable opponent, Eva Louise McNulty. I dedicate this work to her.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 1: Introduction

Chapter 2: Comparing the Mitochondrial Genomes of *Wolbachia*-Dependent and Independent Filarial Nematode Species

Chapter 3: Endosymbiont DNA in Endobacteria-Free Filarial Nematodes Indicates Ancient Horizontal Genetic Transfer

Chapter 4: The Transcriptome and Proteome of a *Wolbachia*-Free Filarial Parasite, *Onchocerca flexuosa*, Provide Evidence of Trans-Domain Horizontal Gene Transfer

Chapter 5: Discussion

Appendix 1: Absence of *Wolbachia* Endobacteria in *Chandlerella quiscali*: an Avian Filarial Parasite

Appendix 2: *Brugia malayi*: Whole Genome Amplification for Genomic Characterization of Filarial Parasites

CHAPTER 1:

Introduction

This chapter was composed entirely by SNM. Comments from Peter U. Fischer and Gary J. Weil were incorporated into the final version.

SECTION I: INTRODUCTION TO FILARIAL NEMATODES

Historical Background

Filarial nematodes are arthropod-borne parasitic worms that infect all classes of vertebrates except fish [1]. The disabling and disfiguring diseases caused by these worms (known collectively as "filariasis") were recognized in ancient times, with potential accounts dating back to ancient Egypt (2000 BCE) and the Nok civilization of West Africa (500 BCE). The Dutch explorer Jan Huygen Linschoten made the first clear description of lymphatic filariasis (LF) in India in the late 16th century, attributing it to a curse upon the descendants of those who martyred St. Thomas [2].

Despite historical knowledge of filarial disease, larval and adult stage parasites were not identified until 1863 and 1876, respectively [2]. The elucidation of the filarial life cycle began in 1877 when Patrick Manson, the father of tropical medicine, found *Wuchereria bancrofti* larvae in mosquitoes that had fed on the blood of his infected gardener [3]. The idea of infection by mosquito bite was confirmed some years later when Manson's assistant observed larval worms in mosquito mouthparts [4]. This is widely considered one of the most significant discoveries in the history of tropical medicine.

The knowledge of the life cycle allowed for some disease prevention through avoidance of the vector and vector control. However, it was not until 1947 that chemotherapeutic agents (first diethylcarbamazine (DEC), later ivermectin and albendazole) were first used to treat human filarial infections [5,6,7]. Despite widespread

and ongoing use of these drugs, filarial diseases continue to be major causes of disability in the developing world.

Life Cycle

Filarial nematodes share a common life cycle (**Fig. 1**), though the species of the vertebrate host and vector may vary from parasite species to parasite species. Following sexual reproduction, females release first-stage larvae (L1) called microfilariae (MF) that reside in the blood or skin of the host. MF are ingested by a blood-feeding arthropod, penetrate the midgut, and develop in muscles or other tissues. Larval worms undergo two molts (from L1 to L2 and from L2 to L3) to reach the infective third larval stage (L3) and migrate to the arthropod's mouthparts for transmission during a subsequent feeding. After infecting the definitive host, worms undergo two more molts (L3 to L4 and L4 to adult) to reach the adult stage. Young adults require further growth and development to reach sexual maturity. The interval between infection and the appearance of MF in the blood or skin (the "prepatent period") varies between species, ranging from two months to more than a year. The life expectancy of adult filarial worms also varies between species, with some surviving more than 10 years.

Phylogeny

Phylogenetic analyses based on the small subunit ribosomal RNA gene suggest that the phylum Nematoda encompasses five distinct clades [8]. Filarial worms belong to clade III (**Fig. 2**). They last shared a common ancestor with the model nematode *Caenorhabditis elegans* (clade V) some 300-500 million years ago [9]. The first filarial

nematodes, likely mosquito-transmitted parasites of crocodilians, are thought to have emerged toward the end of the Jurassic period (150 million years ago) [10]. Diversification within the superfamily Filarioidea most likely occurred during the Tertiary period (65 million – 2.6 million years ago), concurrent with the expansion of the worms' avian and mammalian hosts [11].

The superfamily Filarioidea is comprised of two distinct families, the Filariidae and the Onchocercidae. The Filariidae are divided into two subfamilies (the Filariinae and the Stephanofilariinae) that parasitize various mammals [12]. The Onchocercidae are a more diverse group encompassing approximately 70-80 genera divided into eight subfamilies based on classical parasitological criteria (**Table 1**). Species of this subfamily parasitize a wide array of hosts including amphibians, reptiles, birds and mammals [12].

SECTION II: FILARIAL DISEASE AND TREATMENT

Filarial Disease

Filarial nematodes are well-adapted parasites that infect various hosts throughout the globe. In the developed world, many people are familiar with the heartworm, *Dirofilaria immitis* (Subfamily: Dirofilariinae), which can cause severe disease and death in dogs and cats. According to surveys performed from 2001-2007, approximately 43,500 out of 3 million pet dogs tested in the US were diagnosed with heartworm infection [13]. Human filarial infections mainly occur in the tropics and subtropics. Eight species of filarial nematodes are known to parasitize humans (**Table 2**). Three of the eight,

Onchocerca volvulus, W. bancrofti and *Brugia malayi* (all members of the subfamily Onchocercinae) are responsible for the most serious filarial diseases.

Some 37 million people in 34 countries are infected with *O. volvulus*, the causative agent of African river blindness [14]. Worms of this species are transmitted by blackflies (genus *Simulium*) that breed in oxygen-rich waters of swiftly flowing rivers and streams. Female worms, which can reach 80 cm in length, nest in subcutaneous nodules, or onchocercomata (**Fig. 3A-B**), while the smaller males travel between nodules to mate [15]. MF released from the nodules migrate through skin (**Fig. 3C**), causing dermatitis and alterations in pigmentation and elasticity (**Fig. 3D-E**). Skin disease is often accompanied by discomfort and itching. Inflammatory reactions to MF in the eye can lead to ocular damage or blindness (**Fig. 3F**) [15]. A study in 1995 estimated that approximately 270,000 people had been blinded by *O. volvulus* while another 500,000 were vision impaired, making this parasite the world's second leading infectious cause of blindness [16].

W. bancrofti, *B. malayi* and *Brugia timori* are the causative agents of LF. 120 million people in more than 70 countries are infected with these parasites, and approximately 1.4 billion are at risk [17]. The 2 to 8 cm long adult worms nest in lymphatic vessels and lymph nodes, commonly in the groin or lower extremities. Inflammatory reactions to adult worms damage lymphatic vessels, leading to edema in the limbs (usually the legs) (**Fig. 4A-B**) and hydrocele of the scrotum (**Fig. 4C-D**). This, in turn, leaves the affected regions vulnerable to secondary infections by bacteria and fungi, further exacerbating lymphatic damage and promoting elephantiasis (**Fig. 4E-F**)

[18]. With more than 40 million people suffering significant clinical disease, LF is the world's second leading cause of long-term disability [19].

Unfortunately, no effective cure is available for persons already suffering from the chronic symptoms of LF or river blindness. Elephantiasis patients often experience partial improvement with good hygiene and care of affected limbs, including treatment of secondary infections, but limbs remain swollen. Hydroceles can be corrected surgically, but surgery is rarely available to affected populations (**Fig. 4C-D**). Although onchocercal skin disease improves with the clearance of MF, anthelmintics do not reverse ocular damage caused by the infection.

Combating Filarial Disease in Humans

Due to its tremendous impact on global health and on economic development in endemic countries, the Global Programme to Eliminate Lymphatic Filariasis (GPELF) was implemented in 2000 with the goal of eliminating LF as a public health problem by the year 2020 [20]. Similarly, the African Programme for Onchocerciasis Control, initiated in 1995, seeks to control river blindness and onchocercal skin disease in Africa [14]. Both programs rely on mass drug administration (MDA) to endemic populations. Drugs used for MDA, including DEC, ivermectin and albendazole, are either donated by pharmaceutical companies or available at low cost. Ivermectin is used to combat onchocerciasis, whereas combinations of ivermectin plus albendazole or DEC plus albendazole are used to combat LF. These treatments temporarily sterilize adult worms and eliminate MF, thereby blocking transmission. However, they do not effectively kill the adults, so MDA programs must be sustained over the reproductive lifespan of the

worms in order to eliminate these infections from endemic communities [21,22]. This may take 4-6 years for species that cause LF and 10-15 years for *O. volvulus*.

MDA programs have successfully reduced LF infection rates in many areas. The GPELF provided approximately 1.9 billion MDA treatments to 570 million patients in 48 countries between 2000 and 2007 [23]. Ottesen et al. estimated that these efforts prevented some 3.5 million cases of lymphedema, 6 million cases of hydrocele, and the loss of 32 million disability adjusted life years. The economic benefit of the first 8 years of the GPELF was estimated to have been at least US$21 billion [24]. However, prospects for global elimination of LF and control of onchocerciasis would be much better if drugs were available to efficiently kill or permanently sterilize adult worms. The urgent need to develop adulticidal anti-filarial therapies has inspired interest in alternative drug targets, including *Wolbachia* endobacteria.

SECTION III: *WOLBACHIA* ENDOBACTERIA OF ARTHROPODS AND FILARIAL NEMATODES

The Genus *Wolbachia*

Wolbachia are vertically-transmitted alpha-proteobacteria in the family Rickettsiaceae [25]. *Wolbachia* were first detected in the mosquito *Culex pipiens* in 1924 [26] and have since been found in many species of arthropods (insects, crustaceans, spiders and mites) and filarial nematodes. These bacteria are remarkably well adapted to the intracellular environment where they reside in clusters within host-derived vacuoles.

Infected tissues always include the ovaries and oocytes of female hosts, allowing for vertical transmission from mother to offspring.

Despite considerable genetic and phenotypic variation, the genus *Wolbachia* contains only one recognized species, *Wolbachia pipientis* [27,28]. The various strains of this species, generally denoted by the species of the host (e.g. *Wolbachia* endosymbiont of *B. malayi*), have been subdivided into eight supergroups, designated A-H, based on analyses of various genetic markers (**Fig. 5**) [29,30,31]. Most of the *Wolbachia* strains that infect arthropods cluster into supergroups A and B, while filarial *Wolbachia* cluster into supergroups C and D. Supergroup F is particularly interesting in that it includes strains that infect both insects and filarial hosts [32].

Wolbachia Endobacteria of Arthropods

PCR-based surveys indicate that *Wolbachia* may infect up to 70% of all insect species [33,34]. They typically act as parasites that manipulate host reproduction to promote their own transmission. Phenotypic effects of *Wolbachia* infection are varied and can include: feminization of genetic males**,** parthenogenesis (development of unfertilized eggs), death of male offspring from infected mothers, and cytoplasmic incompatibility [35]. One might expect vertical transmission to select for mutualism rather than parasitism because the fitness of the parasite is inextricably linked to the fitness of the host. However, there are only a few known examples of arthropod *Wolbachia* strains increasing host fitness. Reported positive effects of *Wolbachia* infection include increased fecundity in *Drosophila simulans*, resistance to RNA viruses in *Drosophila melanogaster*, and nutrient provision in the bedbug *Cimex lectularius* [36,37,38].

Wolbachia Endobacteria of Filarial Nematodes

Wolbachia endobacteria were first observed in ultrastructural studies of filarial nematodes in the 1970's [39,40,41]. However, *Wolbachia* were largely overlooked until their rediscovery in the genomic era when bacterial DNA was identified in *B. malayi* cDNA libraries [42,43]. Early surveys of the distribution of *Wolbachia* among filarial nematodes focused mainly on two subfamilies, the Onchocercinae and the Dirofilariinae, as these encompass many of the medically and economically important parasites of humans and domestic animals. The results of these limited studies led to the erroneous conclusion that the vast majority of filarial species were *Wolbachia*-infected and that the five species shown to lack *Wolbachia* (*Acanthocheilonema viteae*, *Loa loa*, *Onchocerca flexuosa*, *Mansonella perstans* and *Setaria equina*) were exceptional [25,44,45,46,47,48,49,50,51,52]. However, recent studies have shown that *Wolbachia* infection is less common in other subfamilies and that filarial parasites of frogs, lizards and birds are likely *Wolbachia*-free [32]. The distribution of *Wolbachia* among filarial nematodes is summarized in **Table 3**.

In adult and juvenile worms of both sexes, the majority of *Wolbachia* reside in the lateral chords, longitudinal expansions of the hypodermis that run the length of the worm body [25]. Recent studies suggest that as juvenile adult worms mature, bacteria migrate from the lateral chords to infect the developing ovaries and testes [53]. Although the oocytes become infected, a process that is vital for vertical transmission of the bacterial infection, spermatozoa do not [53]. See **Figure 6** for a description of the typical distribution of *Wolbachia* in adult worms.

In contrast to the parasitic behavior of insect *Wolbachia*, several lines of evidence suggest that filarial *Wolbachia* are obligate mutualists. First, the phylogenetic relationships between *Wolbachia* strains are nearly identical to the phylogenetic relationships between their filarial hosts [51]. The long, stable co-evolution that led to these matching phylogenies implies interdependence. This differs substantially from the incongruent phylogenetic patterns seen in insects and parasitic *Wolbachia* [35]. Second, with only a few potential exceptions, every individual of an infected filarial species is infected during every life cycle stage [32]. Again, this is in contrast to insect *Wolbachia* where infected and uninfected populations often exist within a given species [35]. Finally and most importantly, studies have shown that long-term (daily for 4-6 weeks) administration of antibiotics effective against *Rickettsia* has a severe impact on *Wolbachia*-dependent filarial worms. For example, Hoerauf et al. administered tetracycline to mice infected with *Litomosoides sigmodontis* (*Wolbachia*-dependent) and gerbils infected with *A. viteae* (*Wolbachia*-independent) [54]. The treatments reduced *Wolbachia* populations in *L. sigmodontis* to levels undetectable by immunohistochemistry and electron microscopy. This decrease in *Wolbachia* was associated with retarded growth and infertility in the worms and reduced worm burdens in the rodent hosts. The same treatment had no effect on *A. viteae*, which are naturally *Wolbachia*-free. Subsequent experiments showed that tetracycline has similar effects on other *Wolbachia*-dependent species [55,56,57].

Antibiotic treatments have also proven efficacious for treating human filariases. Several studies have shown that supplementing traditional treatments (ivermectin for *O. volvulus* or DEC for the lymphatic filariae) with doxycycline/tetracycline results in

enhanced suppression of microfilaremia compared to ivermectin or DEC alone [58,59]. Addition of doxycycline to conventional therapy also enhanced macrofilaricidal effects, killing up to 60% of adult female *O. volvulus* in examined nodules and 77% of adult *W. bancrofti* in lymphatic worm nests [60,61]. Other studies have shown that doxycycline alone is an effective treatment against both LF and onchocerciasis [62,63]. For a comparison of doxycycline to traditional anthelmintics, see **Table 4**.

Doxycycline may be effective for treating filarial infection, but it is not practical for MDA. Ongoing control programs typically treat endemic populations with a single dose of drugs only once or twice per year. However, doxycycline must be administered daily for four to six weeks, after which the MF and adult worm decline is relatively slow [64]. This makes it unsuitable for use in MDA programs. Doxycycline's mode of action may be responsible for the length of treatment required and the slow clearance observed in patient populations. As an inhibitor of bacterial protein translation, it may take some time for this drug to affect specific bacterial pathways or processes required by the worm host. Drugs designed to directly target these critical pathways may work faster. Unfortunately, these pathways have yet to be identified.

Potential Interactions Between *Wolbachia* and Filarial Nematodes

Hypotheses to explain the dependence of filarial worms on *Wolbachia* have arisen from studies that compared the genomes of *B. malayi* and its *Wolbachia* endosymbiont (*wBm*) [65,66]. The presence of a pathway in *wBm* and its absence in *B. malayi* may indicate metabolic provisioning on the part of the bacteria. Of course, one must consider the three-way interaction between the bacteria, the worm and the worm's vertebrate host,

as the worm may scavenge from either partner. However, the pressure to reduce genome size, the impact of which is particularly evident in the small, 1 Mb genome of *wBm* [65], should have eliminated pathways designed to synthesize substances that are readily available from another source. For example, *Wolbachia* have lost the capacity to synthesize most amino acids, NAD, biotin, folate and coenzyme A [65] since these can be obtained from the worm or vertebrate host. Remaining pathways are likely to be required to meet specific needs of the bacteria or needs of both the bacteria and the worm. Specific examples are discussed below.

Unlike the genomes of *Rickettsia* and other intracellular bacteria (i.e. *Buchnera, Blochmannia, Mycoplasma* and *Chlamydia*), *wBm* encodes complete pathways for the *de novo* synthesis of purines and pyrimidines [65]. The fact that *B. malayi* lacks 9 of the 10 enzymes involved in *de novo* purine synthesis (notably, the conversion of phosphoribosyl pyrophosphate to inosine monophosphate) suggests that *Wolbachia* may be synthesizing nucleotides for both organisms [66,67]. Foster et al. theorized that *Wolbachia* metabolites could be particularly important during times of very active cell division (i.e., oogenesis, embryogenesis, molting, etc.) [65]. Coincidentally, tetracycline/doxycycline treatments have their most dramatic effects on worm biology during these times [56,57]. It stands to reason that nucleotides would be vitally important during for intense cell division.

wBm also encodes pathways for the *de novo* synthesis of riboflavin and FAD whereas synthesis pathways for other vitamins and cofactors are missing [65]. This is intriguing because *B. malayi* lacks the 5 enzymes required for riboflavin production [66,67]. Recent findings suggest that *Wolbachia* synthesize B vitamins (including

riboflavin) necessary to sustain growth and fertility in the bedbug *C. lectularius* [36], which makes this pathway particularly interesting.

Another pathway present in *wBm* but absent in *B. malayi* is responsible for C_4-type heme biosynthesis. *wBm* encodes 6 of the 7 enzymes in this pathway [65]. It should be noted that the missing enzyme, protoporphyrinogen-IX oxidase (step 6 in the 7-step pathway), is also absent in other species of heme-producing bacteria [68]; this synthetic step may be carried out by other enzymes. In contrast, *B. malayi* only encodes ferrochelatase, the final enzyme in the heme synthesis pathway [66,67]. Inhibitors of bacterial heme synthesis appear to affect worm motility *in vitro* [69], so *Wolbachia* may be an important source of this vital prosthetic group.

Impact of *Wolbachia* on Host Genetics

Wolbachia have impacted their arthropod and filarial hosts by horizontal gene transfer (HGT). Heritable transfers of DNA from *Wolbachia* to the nuclear genome of the invertebrate host are facilitated by the intimate association between the bacteria and the host germline. This phenomenon was first noted in the bean beetle, *Callosobruchis chinensis* [70]. These beetles were infected with three distinct *Wolbachia* strains. Antibiotic treatments successfully reduced populations of two of these strains to levels undetectable by PCR, but sequences from the third were still readily amplifiable. Further experimentation revealed that the PCR product from the third strain was not derived from lingering endobacteria, but from a large insert of *Wolbachia* DNA nestled in the beetle's X chromosome. A fragment of *Wolbachia* DNA was also discovered in *O. volvulus* and the closely related cattle parasite *Onchocerca ochengi* [71]. The identification of the

same insert in both species led the authors to hypothesize that the transfer had occurred in a common ancestor. Other studies in insects and filarial nematodes indicated that HGT from *Wolbachia* may be a common occurrence. Dunning-Hotopp et al. identified *Wolbachia* sequences in several insect and *Wolbachia*-dependent filarial nematode species [72]. This included nearly 1 Mb of *Wolbachia* DNA in the nuclear genome of *Drosophila ananassae* and 249 *B. malayi* genomic contigs with *Wolbachia* inserts of various sizes. RT-PCR studies showed RNA-level expression of some of the *Wolbachia* genes in antibiotic-cured *D. ananassae*. Unfortunately, the expression of transferred genes cannot be proven in *Wolbachia*-dependent filarial nematodes because there is no way to determine whether transcripts of interest are derived from the worm genome or from the endobacteria since the *Wolbachia* cannot be eliminated.

HGT is not the only means by which *Wolbachia* affect host genomes. Mitochondria are particularly sensitive to the evolutionary pressure exerted by *Wolbachia*. Due to their reproductive effects, *Wolbachia* infections often lead to bottlenecks in insect populations, as the infected insects outbreed the uninfected ones. *Wolbachia* and mitochondria are co-inherited from the mother worm, so the expansion of the *Wolbachia*-infected population is concurrent with the expansion of the mitochondrial haplotypes associated with these individuals. These *Wolbachia*-mitochondria "sweeps," characterized by unusually low degrees of mitochondrial variation in infected populations, have been noted in many insect species [73,74,75,76,77]. A similar lack of mitochondrial diversity has been reported in populations of *D. immitis* in comparison to *Wolbachia*-uninfected non-filarial nematodes [78]. These studies raise questions

regarding the effect(s) of *Wolbachia* on mitochondrial genome content and organization in filarial nematodes.

SECTION IV: RESOURCES AND CHALLENGES IN THE STUDY OF FILARIAL NEMATODES

Resources

Although the availability of parasite material can be a limiting factor in the study of filarial nematodes, material from several species is readily available. A few species, like *B. malayi* (*Wolbachia*-dependent) and *A. viteae* (*Wolbachia*-dependent), can be maintained in the laboratory in convenient small animal models [79,80]. Others, like the deer parasite *O. flexuosa*, can be collected from hunted animals [50]. *Chandlerella quiscali*, a *Wolbachia*-free filarial nematode that parasitizes the brain of the common grackle, can be collected throughout most of the United States and Canada (see **Appendix 1**). Ongoing international projects allow for the collection of limited amounts of *O. volvulus* worms and *W. bancrofti* and *L. loa* microfilariae from infected humans. In these cases, whole genome amplification (see **Appendix 2**) can be used to produce sufficient quantities of DNA for experimentation.

Recent years have seen a tremendous increase in genomic studies of filarial nematodes and related organisms. The 95 Mb genome of *B. malayi* has been sequenced to an estimated 9X coverage [66,67]. Although technically incomplete, it is assumed that the majority of *B. malayi* genes have been accounted for. Other species involved in the *B. malayi* life cycle, including the human host [81,82], mosquito vectors (*Aedes* and

Anopheles) [83,84], and the *Wolbachia* endosymbiont [65] have also been sequenced, as have various strains of insect *Wolbachia* [85,86,87]. The genomes and transcriptomes sequences of other filarial species such as *L. loa*, *O. volvulus* and *W. bancrofti* are in progress.

Although *C. elegans* is a very distant relative of filarial nematodes, it is still an important model for comparative studies. At the structural level, most nematodes have similar body plans and a similar pattern of development. At the genomic level, approximately 50% of all *B. malayi* genes have clear orthologs in *C. elegans* [67]. The wealth of genomic, transcriptomic, and proteomic data available for *C. elegans* can be used, with appropriate caution, to make inferences regarding the function and behavior of these genes. Other resources and reagents developed for use with *C. elegans* can sometimes be applied in the study of filarial nematodes. For example, unpublished studies from the Weil lab have shown that many monoclonal antibodies to *C. elegans* proteins also bind targets in *B. malayi*; many of these can be used in immunohistochemistry.

Obstacles to the Study of Filarial Nematodes and *Wolbachia*

One of the major obstacles impeding the study of parasitic nematodes is the lack of *in vitro* culture systems or convenient animal hosts to maintain many species in the laboratory. *O. volvulus* and *W. bancrofti* serve as notable examples, as material from these species must be collected from humans in endemic regions. Nodules containing adult *O. volvulus* can be surgically recovered from consenting patients. However, nodulectomies are not performed as part of medical treatment unless the nodules are on

or near the face or otherwise interfere with daily activities. In the case of *W. bancrofti*, microfilaria can be filtered from infected blood, but the adult worms nestled in lymphatic vessels are rarely recovered. The same goes for *B. timori* and anthrophilic strains of *B. malayi*. In central Africa, patients infected with *Loa loa* and *Mansonella* are often co-infected with other filarial species, which complicates the study of these parasites.

Another challenge in the study of filarial nematodes is genetic intractability. Stable transgenesis and methods for gene deletion are not yet available for filarial parasites. RNAi has been used to knock down expression of several genes in *B. malayi*, *O. volvulus* and *L. sigmodontis* [88,89,90], but results are inconsistent. Skepticism is fueled by the fact that genes encoding major components of the RNAi apparatus appear to be missing from the *B. malayi* genome [66]. A new RNAi protocol has shown success in mosquito-infecting larval parasites [91], but this may be facilitated by RNAi machinery present in the mosquito rather than the worm.

Research on filarial nematodes is further hindered by a lack of genetic information for the vast majority of the species identified to date. This is particularly true for *Wolbachia*-independent filarial species. With the exception of *L. loa*, these are mainly animal parasites of little economic importance. Therefore, there is little incentive to sequence and annotate their genomes. Comparative studies of closely and/or distantly related species could help to answer several important biological questions, but these studies will not be possible until more genomic information becomes available.

SECTION V: SCOPE AND AIMS OF THE THESIS

One looming question in the field of filariasis research deals with the relationship between these organisms and their *Wolbachia* endobacteria. Why is it that some species are fully dependent on *Wolbachia* for proper development, fertility and survival, while other species thrive in the absence of a bacterial partner? Answering this question would not only be interesting from a biological and evolutionary point of view, but it could also have profound practical importance. *Wolbachia* components or metabolic products that are required for nematode survival could serve as useful targets for novel, potentially superior, antifilarial chemotherapies.

I proposed two separate but not mutually exclusive hypotheses to explain *Wolbachia* dependence in filarial nematodes. The first hypothesis reasoned that this dependence resulted from a loss of gene function in *Wolbachia*-infected worms. Endosymbiont provision of essential materials could remove selective pressure for the maintenance of redundant or unnecessary genes in the filarial worm, and this could lead to gene loss due to mutation or deletion. To test this hypothesis, I chose to focus on the mitochondrial genome. *Wolbachia* are closely related to the theoretical mitochondrial progenitor [65,92], and they could provide a similar function (e.g., energy production). Mitochondria are also particularly sensitive to evolutionary pressures exerted by *Wolbachia* infection, as was previously discussed. We reasoned that the mitochondrial genomes of *Wolbachia*-dependent filarial nematodes might be deficient compared to those of *Wolbachia*-independent species. For instance, genes related to mitochondrial respiration might be missing or altered in expression or function. In such cases, the

endobacterial partner may augment the mitochondrial genome in a beneficial way with its own complement of genes. Chapter 2 of this thesis deals with the sequencing and analysis of the mitochondrial genomes of several *Wolbachia*-dependent and independent filarial nematode species. We sought to determine how mitochondrial genomes from *Wolbachia*-dependent and independent species compare with respect to content, gene order and sequence.

My second hypothesis proposed that filarial species that presently lack *Wolbachia* may have been *Wolbachia*-dependent in the distant past and that horizontal transfer of critical genes from the endosymbiont into the nuclear genome of the nematode host may have rendered the bacteria expendable. As previously discussed, *Wolbachia* DNA inserts have been identified in several species of insects and *Wolbachia*-dependent filarial nematodes. Chapters 3 and 4 of this thesis describe the search for evidence of horizontal gene transfer from *Wolbachia* to filarial species that no longer contain an endosymbiont. Following the identification of *Wolbachia*-like DNA sequences in the nuclear genomes of these species, we looked for evidence of expression of these genes at the RNA and protein levels as a first step in determining the impact of horizontally transferred genes and gene fragments on the survival of *Wolbachia*-independent filarial species.

REFERENCES

1. Anderson RC, Bain O (1976) Keys to genera of the order Spirurida. Part 3. Diplotriaenoidea, Aproctoidea and Filarioidea. In: Anderson RC, Chabaud AG, Willmott S, editors. CIH keys to the nematode parasites of vertebrates. Farnham Royal, UK: Commonwealth Agricultural Bureau. pp. 59-116.
2. Cox FE (2002) History of human parasitology. Clin Microbiol Rev 15: 595-612.
3. Manson P, Cobbold FKS (1878) On the development of *Filaria sanguinis hominis*, and on the mosquito considered as a nurse. Journal of the Linnean Society of London, Zoology 14: 304-311.
4. Low GC (1900) A Recent Observation on Filaria Nocturna in *Culex*: Probable Mode of Infection of Man. Br Med J 1: 1456-1457.
5. Santiago-Stevenson D, Oliver-Gonzalez J, Hewitt R (1947) Treatment of *Filariasis bancrofti* with 1-diethylcarbamyl-4-methylpiperazine hydrochloride (hetrazan). J Am Med Assoc 135: 708-712.
6. Aziz MA, Diallo S, Diop IM, Lariviere M, Porta M (1982) Efficacy and tolerance of ivermectin in human onchocerciasis. Lancet 2: 171-173.
7. Ottesen EA, Duke BO, Karam M, Behbehani K (1997) Strategies and tools for the control/elimination of lymphatic filariasis. Bull World Health Organ 75: 491-503.
8. Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, et al. (1998) A molecular evolutionary framework for the phylum Nematoda. Nature 392: 71-75.
9. Vanfleteren JR, Van de Peer Y, Blaxter ML, Tweedie SA, Trotman C, et al. (1994) Molecular genealogy of some nematode taxa as based on cytochrome c and globin amino acid sequences. Mol Phylogenet Evol 3: 92-101.
10. Chabaud AG, Bain O (1994) The evolutionary expansion of the Spirurida. Int J Parasitol 24: 1179-1201.
11. Bain O (2002) Evolutionary Relationships Among Filarial Nematodes. In: Klei TR, Rajan TV, editors. World Class Parasites: Volume 5 The Filaria. Boston: Kluwer Academic Publishers. pp. 21-29.
12. Anderson RC (2000) Nematode Parasites of Vertebrates: Their Development and Transmission. New York, New York, USA: CABI Publishing.
13. Bowman D, Little SE, Lorentzen L, Shields J, Sullivan MP, et al. (2009) Prevalence and geographic distribution of *Dirofilaria immitis*, *Borrelia burgdorferi*, *Ehrlichia canis*, and *Anaplasma phagocytophilum* in dogs in the United States: results of a national clinic-based serologic survey. Vet Parasitol 160: 138-148.
14. Remme JHF, Feenstra P, Lever PR, Medici AC, Morel CM, et al. (2006) Tropical Diseases Targeted for Elimination: Chagas Disease, Lymphatic Filariasis, Onchocerciasis, and Leprosy. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M et al., editors. Disease Control Priorities in Developing Countries. 2 ed. Washington DC: World Bank.
15. Brattig NW (2004) Pathogenesis and host responses in human onchocerciasis: impact of *Onchocerca* filariae and *Wolbachia* endobacteria. Microbes Infect 6: 113-128.
16. WHO (1995) Onchocerciasis and its control. Report of a WHO Expert Committee on Onchocerciasis Control. World Health Organ Tech Rep Ser 852: 1-104.
17. WHO (2009) Global programme to eliminate lymphatic filariasis. Wkly Epidemiol Rec 84: 437-444.

18. Olszewski WL, Jamal S, Manokaran G, Pani S, Kumaraswami V, et al. (1997) Bacteriologic studies of skin, tissue fluid, lymph, and lymph nodes in patients with filarial lymphedema. Am J Trop Med Hyg 57: 7-15.
19. Gyapong JO, Kumaraswami V, Biswas G, Ottesen EA (2005) Treatment strategies underpinning the global programme to eliminate lymphatic filariasis. Expert Opin Pharmacother 6: 179-200.
20. Ottesen EA (2000) The global programme to eliminate lymphatic filariasis. Trop Med Int Health 5: 591-594.
21. Klager SL, Whitworth JA, Downham MD (1996) Viability and fertility of adult *Onchocerca volvulus* after 6 years of treatment with ivermectin. Trop Med Int Health 1: 581-589.
22. Weil GJ, Kastens W, Susapu M, Laney SJ, Williams SA, et al. (2008) The impact of repeated rounds of mass drug administration with diethylcarbamazine plus albendazole on bancroftian filariasis in Papua New Guinea. PLoS Negl Trop Dis 2: e344.
23. Ottesen EA, Hooper PJ, Bradley M, Biswas G (2008) The global programme to eliminate lymphatic filariasis: health impact after 8 years. PLoS Negl Trop Dis 2: e317.
24. Chu BK, Hooper PJ, Bradley MH, McFarland DA, Ottesen EA (2010) The economic benefits resulting from the first 8 years of the Global Programme to Eliminate Lymphatic Filariasis (2000-2007). PLoS Negl Trop Dis 4: e708.
25. Taylor MJ, Bandi C, Hoerauf A (2005) *Wolbachia* bacterial endosymbionts of filarial nematodes. Adv Parasitol 60: 245-284.
26. Hertig M, Wolbach SB (1924) Studies on *Rickettsia*-Like Micro-Organisms in Insects. J Med Res 44: 329-374 327.
27. Dumler JS, Barbet AF, Bekker CP, Dasch GA, Palmer GH, et al. (2001) Reorganization of genera in the families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int J Syst Evol Microbiol 51: 2145-2165.
28. Lo N, Paraskevopoulos C, Bourtzis K, O'Neill SL, Werren JH, et al. (2007) Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. Int J Syst Evol Microbiol 57: 654-657.
29. Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, et al. (2006) Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Appl Environ Microbiol 72: 7098-7110.
30. Bordenstein SR, Paraskevopoulos C, Dunning Hotopp JC, Sapountzis P, Lo N, et al. (2009) Parasitism and mutualism in *Wolbachia*: what the phylogenomic trees can and cannot say. Mol Biol Evol 26: 231-241.
31. Casiraghi M, Bordenstein SR, Baldo L, Lo N, Beninati T, et al. (2005) Phylogeny of *Wolbachia pipientis* based on gltA, groEL and ftsZ gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. Microbiology 151: 4015-4022.

32. Ferri E, Bain O, Barbuto M, Martin C, Lo N, et al. (2011) New Insights into the Evolution of *Wolbachia* Infections in Filarial Nematodes Inferred from a Large Range of Screened Species. PLoS One 6: e20843.
33. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008) How many species are infected with *Wolbachia*?--A statistical analysis of current data. FEMS Microbiol Lett 281: 215-220.
34. Jeyaprakash A, Hoy MA (2000) Long PCR improves *Wolbachia* DNA amplification: wsp sequences found in 76% of sixty-three arthropod species. Insect Mol Biol 9: 393-405.
35. Werren JH, Baldo L, Clark ME (2008) *Wolbachia*: master manipulators of invertebrate biology. Nat Rev Microbiol 6: 741-751.
36. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T (2010) *Wolbachia* as a bacteriocyte-associated nutritional mutualist. Proc Natl Acad Sci U S A 107: 769-774.
37. Teixeira L, Ferreira A, Ashburner M (2008) The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. PLoS Biol 6: e2.
38. Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA (2007) From parasite to mutualist: rapid evolution of *Wolbachia* in natural populations of Drosophila. PLoS Biol 5: e114.
39. Kozek WJ (1977) Transovarially-transmitted intracellular microorganisms in adult and larval stages of *Brugia malayi*. J Parasitol 63: 992-1000.
40. Kozek WJ, Marroquin HF (1977) Intracytoplasmic bacteria in *Onchocerca volvulus*. Am J Trop Med Hyg 26: 663-678.
41. McLaren DJ, Worms MJ, Laurence BR, Simpson MG (1975) Micro-organisms in filarial larvae (Nematoda). Trans R Soc Trop Med Hyg 69: 509-514.
42. Henkle-Duhrsen K, Eckelt VH, Wildenburg G, Blaxter M, Walter RD (1998) Gene structure, activity and localization of a catalase from intracellular bacteria in *Onchocerca volvulus*. Mol Biochem Parasitol 96: 69-81.
43. Williams SA, Lizotte-Waniewski MR, Foster J, Guiliano D, Daub J, et al. (2000) The filarial genome project: analysis of the nuclear, mitochondrial and endosymbiont genomes of *Brugia malayi*. Int J Parasitol 30: 411-419.
44. Bandi C, Trees AJ, Brattig NW (2001) *Wolbachia* in filarial nematodes: evolutionary aspects and implications for the pathogenesis and treatment of filarial diseases. Vet Parasitol 98: 215-238.
45. Taylor MJ, Hoerauf A (1999) *Wolbachia* bacteria of filarial nematodes. Parasitol Today 15: 437-442.
46. Bandi C, Anderson TJ, Genchi C, Blaxter ML (1998) Phylogeny of *Wolbachia* in filarial nematodes. Proc Biol Sci 265: 2407-2413.
47. Buttner DW, Wanji S, Bazzocchi C, Bain O, Fischer P (2003) Obligatory symbiotic *Wolbachia* endobacteria are absent from *Loa loa*. Filaria J 2: 10.
48. Chirgwin SR, Porthouse KH, Nowling JM, Klei TR (2002) The filarial endosymbiont *Wolbachia* sp. is absent from *Setaria equina*. J Parasitol 88: 1248-1250.
49. McGarry HF, Pfarr K, Egerton G, Hoerauf A, Akue JP, et al. (2003) Evidence against *Wolbachia* symbiosis in *Loa loa*. Filaria J 2: 9.

50. Plenge-Bonig A, Kromer M, Buttner DW (1995) Light and electron microscopy studies on *Onchocerca jakutensis* and *O. flexuosa* of red deer show different host-parasite interactions. Parasitol Res 81: 66-73.
51. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. Int J Parasitol 34: 191-203.
52. Grobusch MP, Kombila M, Autenrieth I, Mehlhorn H, Kremsner PG (2003) No evidence of *Wolbachia* endosymbiosis with *Loa loa* and *Mansonella perstans*. Parasitol Res 90: 405-408.
53. Fischer K, Beatty WL, Jiang D, Weil GJ, Fischer PU (2011) Tissue and Stage-Specific Distribution of *Wolbachia* in *Brugia malayi*. PLoS Negl Trop Dis 5: e1174.
54. Hoerauf A, Nissen-Pahle K, Schmetz C, Henkle-Duhrsen K, Blaxter ML, et al. (1999) Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility. J Clin Invest 103: 11-18.
55. Bandi C, McCall JW, Genchi C, Corona S, Venco L, et al. (1999) Effects of tetracycline on the filarial worms *Brugia pahangi* and *Dirofilaria immitis* and their bacterial endosymbionts *Wolbachia*. Int J Parasitol 29: 357-364.
56. Casiraghi M, McCall JW, Simoncini L, Kramer LH, Sacchi L, et al. (2002) Tetracycline treatment and sex-ratio distortion: a role for *Wolbachia* in the moulting of filarial nematodes? Int J Parasitol 32: 1457-1468.
57. Chirgwin SR, Nowling JM, Coleman SU, Klei TR (2003) *Brugia pahangi* and *Wolbachia*: the kinetics of bacteria elimination, worm viability, and host responses following tetracycline treatment. Exp Parasitol 103: 16-26.
58. Hoerauf A, Mand S, Adjei O, Fleischer B, Buttner DW (2001) Depletion of *Wolbachia* endobacteria in *Onchocerca volvulus* by doxycycline and microfilaridermia after ivermectin treatment. Lancet 357: 1415-1416.
59. Supali T, Djuardi Y, Pfarr KM, Wibowo H, Taylor MJ, et al. (2008) Doxycycline treatment of *Brugia malayi*-infected persons reduces microfilaremia and adverse reactions after diethylcarbamazine and albendazole treatment. Clin Infect Dis 46: 1385-1393.
60. Hoerauf A, Specht S, Buttner M, Pfarr K, Mand S, et al. (2008) *Wolbachia* endobacteria depletion by doxycycline as antifilarial therapy has macrofilaricidal activity in onchocerciasis: a randomized placebo-controlled study. Med Microbiol Immunol 197: 295-311.
61. Mand S, Pfarr K, Sahoo PK, Satapathy AK, Specht S, et al. (2009) Macrofilaricidal activity and amelioration of lymphatic pathology in bancroftian filariasis after 3 weeks of doxycycline followed by single-dose diethylcarbamazine. Am J Trop Med Hyg 81: 702-711.
62. Hoerauf A, Mand S, Volkmann L, Buttner M, Marfo-Debrekyei Y, et al. (2003) Doxycycline in the treatment of human onchocerciasis: Kinetics of *Wolbachia* endobacteria reduction and of inhibition of embryogenesis in female *Onchocerca* worms. Microbes Infect 5: 261-273.
63. Turner JD, Tendongfor N, Esum M, Johnston KL, Langley RS, et al. (2010) Macrofilaricidal activity after doxycycline only treatment of *Onchocerca volvulus*

in an area of *Loa loa* co-endemicity: a randomized controlled trial. PLoS Negl Trop Dis 4: e660.
64. Taylor MJ, Hoerauf A, Bockarie M (2010) Lymphatic filariasis and onchocerciasis. Lancet 376: 1175-1185.
65. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, et al. (2005) The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. PLoS Biol 3: e121.
66. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al. (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317: 1756-1760.
67. Scott AL, Ghedin E (2009) The genome of *Brugia malayi* - all worms are not created equal. Parasitol Int 58: 6-11.
68. Panek H, O'Brian MR (2002) A whole genome view of prokaryotic haem biosynthesis. Microbiology 148: 2273-2282.
69. Wu B, Novelli J, Foster J, Vaisvila R, Conway L, et al. (2009) The heme biosynthetic pathway of the obligate *Wolbachia* endosymbiont of *Brugia malayi* as a potential anti-filarial drug target. PLoS Negl Trop Dis 3: e475.
70. Kondo N, Nikoh N, Ijichi N, Shimada M, Fukatsu T (2002) Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci U S A 99: 14280-14285.
71. Fenn K, Conlon C, Jones M, Quail MA, Holroyd NE, et al. (2006) Phylogenetic relationships of the *Wolbachia* of nematodes and arthropods. PLoS Pathog 2: e94.
72. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753-1756.
73. Delgado AM, Cook JM (2009) Effects of a sex-ratio distorting endosymbiont on mtDNA variation in a global insect pest. BMC Evol Biol 9: 49.
74. Hurst GD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc Biol Sci 272: 1525-1534.
75. Raychoudhury R, Grillenberger BK, Gadau J, Bijlsma R, van de Zande L, et al. (2010) Phylogeography of *Nasonia vitripennis* (Hymenoptera) indicates a mitochondrial-*Wolbachia* sweep in North America. Heredity 104: 318-326.
76. Turelli M, Hoffmann AA (1991) Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature 353: 440-442.
77. Turelli M, Hoffmann AA, McKechnie SW (1992) Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. Genetics 132: 713-723.
78. Belanger DH, Perkins SL (2010) *Wolbachia* infection and mitochondrial diversity in the canine heartworm (*Dirofilaria immitis*). Mitochondrial DNA 21: 227-233.
79. Ash LR, Riley JM (1970) Development of subperiodic *Brugia malayi* in the jird, *Meriones unguiculatus*, with notes on infections in other rodents. J Parasitol 56: 969-973.
80. Lucius R, Textor G (1995) *Acanthocheilonema viteae*: rational design of the life cycle to increase production of parasite material using less experimental animals. Appl Parasitol 36: 22-33.

81. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860-921.
82. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. (2001) The sequence of the human genome. Science 291: 1304-1351.
83. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298: 129-149.
84. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, et al. (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316: 1718-1723.
85. Klasson L, Walker T, Sebaihia M, Sanders MJ, Quail MA, et al. (2008) Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. Mol Biol Evol 25: 1877-1887.
86. Klasson L, Westberg J, Sapountzis P, Naslund K, Lutnaes Y, et al. (2009) The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. Proc Natl Acad Sci U S A 106: 5725-5730.
87. Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, et al. (2004) Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. PLoS Biol 2: E69.
88. Aboobaker AA, Blaxter ML (2003) Use of RNA interference to investigate gene function in the human filarial nematode parasite *Brugia malayi*. Mol Biochem Parasitol 129: 41-51.
89. Lustigman S, Zhang J, Liu J, Oksov Y, Hashmi S (2004) RNA interference targeting cathepsin L and Z-like cysteine proteases of *Onchocerca volvulus* confirmed their essential function during L3 molting. Mol Biochem Parasitol 138: 165-170.
90. Pfarr KM, Heider U, Schmetz C, Buttner DW, Hoerauf A (2008) The mitochondrial heat shock protein 60 (HSP60) is up-regulated in *Onchocerca volvulus* after the depletion of *Wolbachia*. Parasitology 135: 529-538.
91. Song C, Gallup JM, Day TA, Bartholomay LC, Kimber MJ (2010) Development of an in vivo RNAi protocol to investigate gene function in the filarial nematode, *Brugia malayi*. PLoS Pathog 6: e1001239.
92. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, et al. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396: 133-140.
93. Gilleard JS (2004) The use of *Caenorhabditis elegans* in parasitic nematode research. Parasitology 128 Suppl 1: S49-70.
94. Morales-Hojas R (2009) Molecular systematics of filarial parasites, with an emphasis on groups of medical and veterinary importance, and its relevance for epidemiology. Infect Genet Evol 9: 748-759.
95. Nanduri J, Kazura JW (1989) Clinical and laboratory aspects of filariasis. Clin Microbiol Rev 2: 39-50.
96. Vincent AL, Ash LR, Frommes SP (1975) The ultrastructure of adult *Brugia malayi* (Brug, 1927) (Nematoda: Filarioidea). J Parasitol 61: 499-512.
97. Junker K, Barbuto M, Casiraghi M, Martin C, Uni S, et al. (2009) *Litomosa chiropterorum* Ortlepp, 1932 (Nematoda: Filarioidea) from a South African miniopterid: redescription, *Wolbachia* screening and phylogenetic relationships with *Litomosoides*. Parasite 16: 43-50.

98. Keiser PB, Coulibaly Y, Kubofcik J, Diallo AA, Klion AD, et al. (2008) Molecular identification of *Wolbachia* from the filarial nematode *Mansonella perstans*. Mol Biochem Parasitol 160: 123-128.
99. Determann A, Mehlhorn H, Ghaffar FA (1997) Electron microscope observations on *Onchocerca ochengi* and *O. fasciata* (Nematoda:Filarioidea). Parasitol Res 83: 591-603.
100. McNulty SN, Fischer K, Mehus JO, Vaughan JA, Tkach VV, et al. (2011) Absence of *Wolbachia* Endobacteria in *Chandlerella quiscali*: An Avian Filarial Parasite. J Parasitol.
101. Basanez MG, Pion SD, Boakes E, Filipe JA, Churcher TS, et al. (2008) Effect of single-dose ivermectin on *Onchocerca volvulus*: a systematic review and meta-analysis. Lancet Infect Dis 8: 310-322.
102. Fox LM (2006) Ivermectin: uses and impact 20 years on. Curr Opin Infect Dis 19: 588-593.
103. Hoerauf A (2008) Filariasis: new drugs and new opportunities for lymphatic filariasis and onchocerciasis. Curr Opin Infect Dis 21: 673-681.
104. Osei-Atweneboana MY, Awadzi K, Attah SK, Boakye DA, Gyapong JO, et al. (2011) Phenotypic evidence of emerging ivermectin resistance in *Onchocerca volvulus*. PLoS Negl Trop Dis 5: e998.

FIGURES

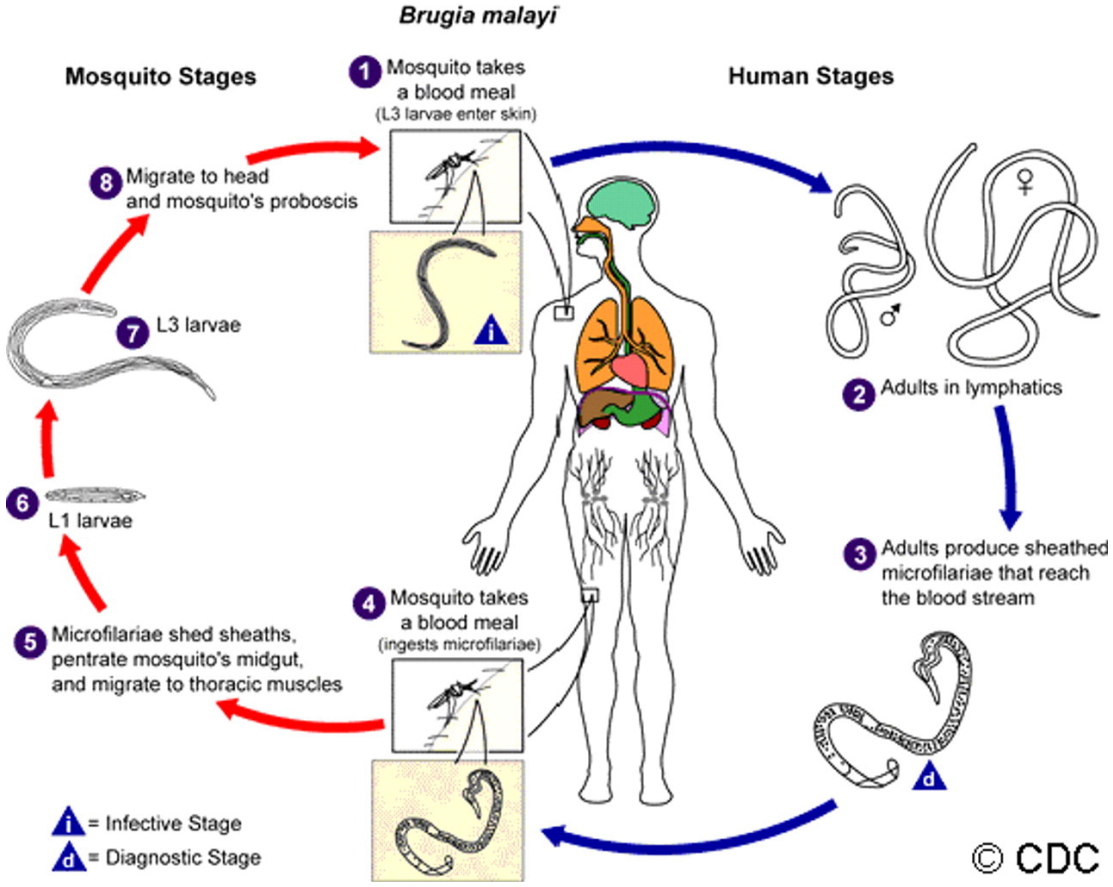


Figure 1: The life cycle of *Brugia malayi* (Family: Onchocercidae, Subfamily: Onchocercinae) is representative of all filarial nematodes. This figure was taken from the website of the Centers for Disease Control and Prevention (http://www.cdc.gov/parasites/lymphaticfilariasis/biology_b_malayi. html).

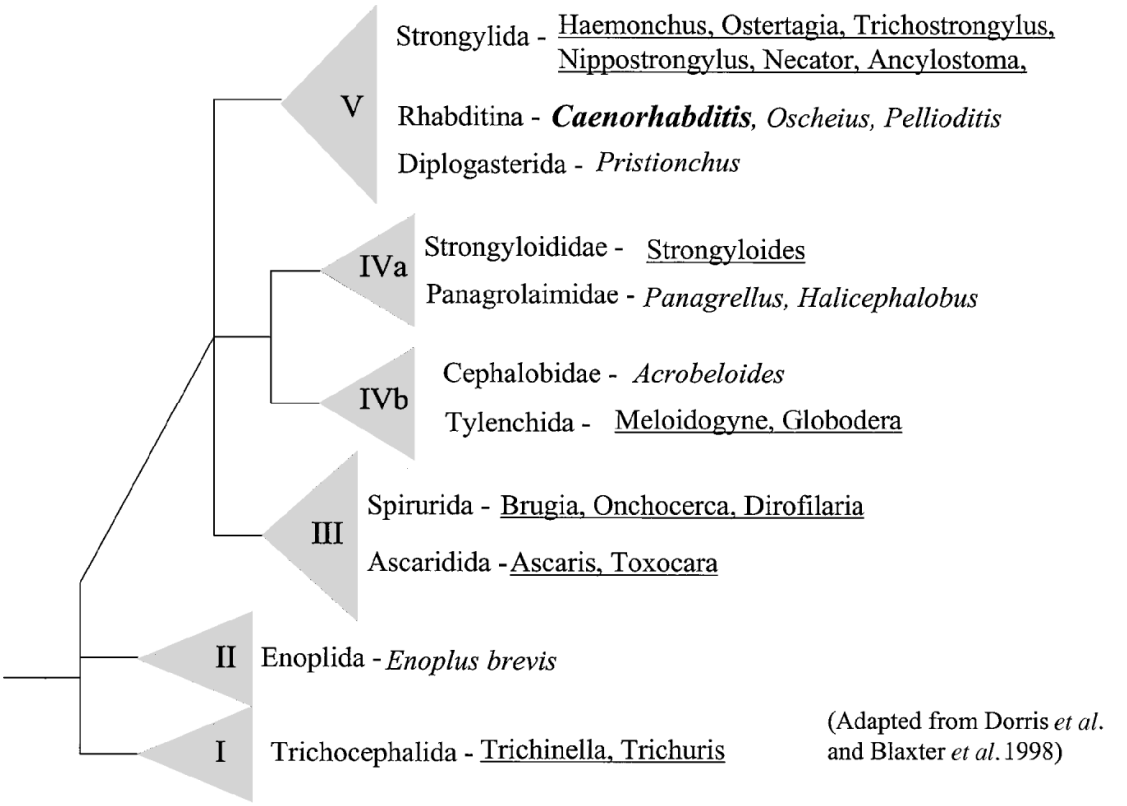


Figure 2: Phylogenetic organization of the phylum Nematoda based on small subunit ribosomal RNA gene sequences. Underlined genera are parasitic whereas italicized genera are free-living. The filarial nematodes are located within clade III, while the free-living model nematode *Caenorhabditis elegans* is a member of clade V. This figure was taken from Gilleard, 2004 [93].

Figure 3: Images of onchocerciasis. *Onchocerca* nodules can often be seen on the trunks of infected patients (A). Nodules can be surgically removed and opened to reveal adult female worms (B). Pathology related to *O. volvulus* infection is caused by microfilaria (see arrows) migrating through the dermis (C). The immune response to circulating microfilaria can result in the severe dermatitis and hyperpigmentation characteristic of sowda (D) or the loss of pigmentation characteristic of "leopard skin" (E). When microfilaria enter the eye, damage to the retina can result in river blindness (F). Photos were provided by D.W. Buttner.

Figure 4: Images of lymphatic filariasis. Blockage of the lymphatic vessels in the groin can lead to mild (A) or severe (B) lymphedema, usually in the lower extremities, or hydrocele of the scrotum (C). Unlike lymphedema of the limbs, a hydrocele can be surgically drained, but this may only provide partial relief (D). Lymphedema can lead to secondary skin changes (E-F). Photos were provided by D.W. Buttner and Peter U. Fischer.

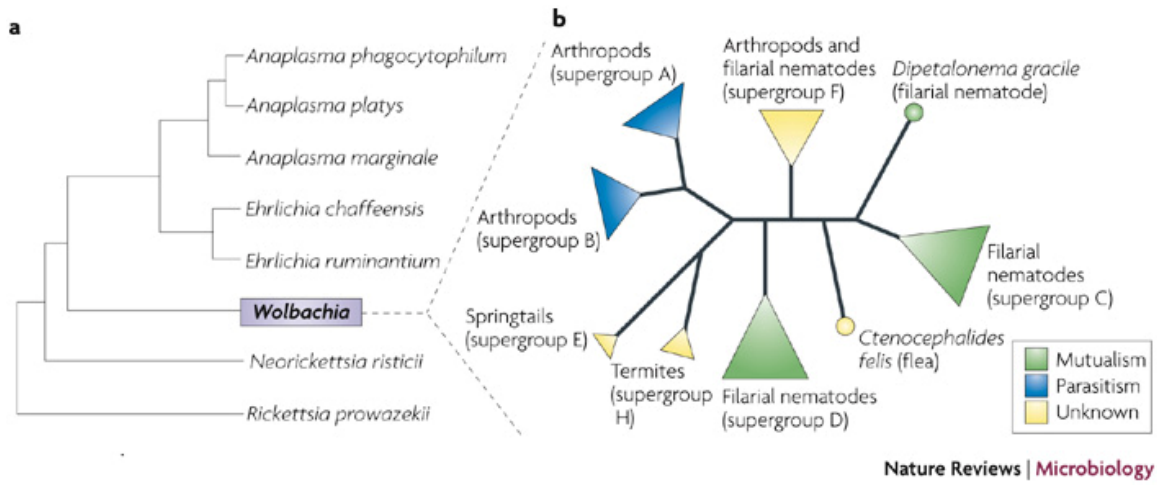


Figure 5: The supergroups of *Wolbachia pipientis*. The genus *Wolbachia* belongs to the *Anaplasmataceae* family within the order Rickettsiales (a). An unrooted phylogenetic tree highlights the main supergroups of the *Wolbachia* genus (b). This figure was taken from Werren et al., 2008 [35].

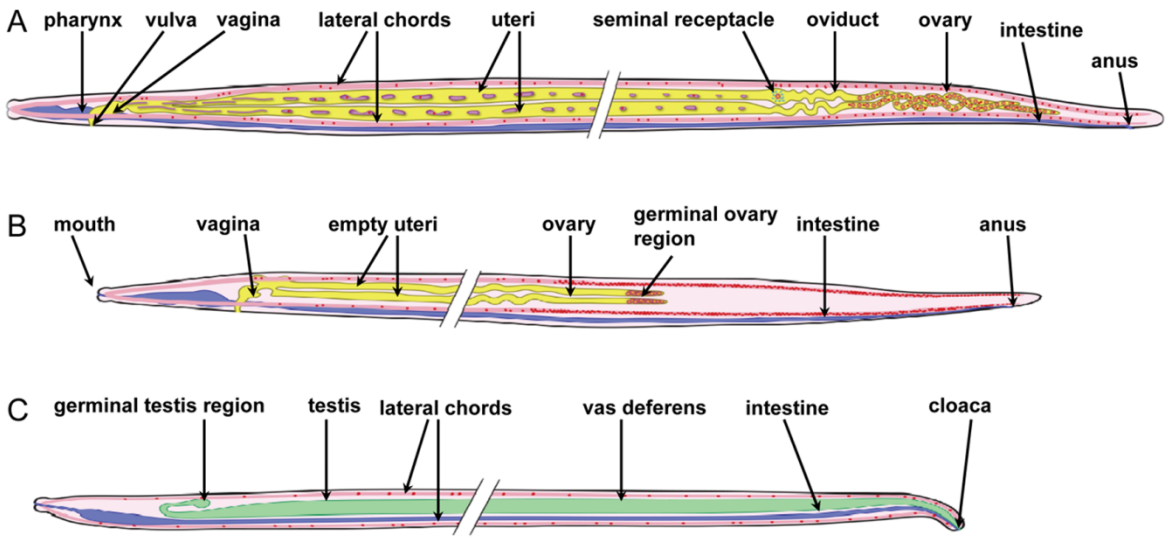


Figure 6: The distribution of *Wolbachia* endobacteria in adult filarial worms. Longitudinal cross-sections of female (**A** and **B**) and male (**C**) worms are depicted from a top (dorsal) viewpoint. In mature adult females, *Wolbachia* (red dots) are present in the lateral chords, ovaries, and in developing embryos that fill the uteri (**A**). In immature adult female worms, populations of bacteria from the lateral chords invade the ovaries as they grow towards the posterior end of the worm (**B**). In males, *Wolbachia* mainly reside in the lateral chords with remnant populations in the *vas deferens*, but not in spermatids or spermatozoa (**C**). This figure was taken from Fischer et al., 2011 [53].

TABLES

Table 1: Classification of the Onchocercinae. Information was taken from Anderson, 2000 [12]. An asterisk indicates representation by at least one sequence in GenBank.

Subfamily	**Host(s)**	**Example Genera**
Oswaldofilariinae	Lacertilians Crocodilians	*Conspiculum* *Oswaldofilaria*
Icosiellinae	Amphibians	*Icosiella*
Waltonellinae	Frogs Toads	*Waltonella**
Setariinae	Artiodactyls Hyracoids Equines	*Setaria**
Dirofilariinae	Mammals Reptiles Birds	*Dirofilaria** *Foleyella** *Loa** *Loaina* *Pelecitus*
Onchocercinae	Mammals Reptiles	*Acanthocheilonema** *Breinlia* *Brugia** *Cercopithifilaria** *Cherylia* *Deraiophoronema** *Dipetalonema** *Elaeophora* *Litomosoides** *Macdonaldius* *Mansonella** *Molinema* *Monanema* *Onchocerca** *Skrjabinofilaria* *Wuchereria** *Yatesia*
Splendidofilariinae	Reptiles Mammals Birds	*Aproctella* *Cardiofilaria* *Chandlerella** *Splendidofilaria* *Thamugadia*
Lemdaninae	Birds	*Eufilaria* *Eulimdana* *Sarconema** *Saurositus*

Table 2: Filarial parasites of humans. This information was collected from Nanduri et al., 1989 and Morales-Hojas, 2009 [94,95].

Species	**Distribution**	**Vectors**	**Location of Adults**	**Location of Microfilaria**	**Common Disease Symptoms**	***Wolbachia* Status**
Wuchereria bancrofti	tropical climates	Culicine and Anopheline mosquitoes	lymphatics	blood	lymphangitis, elephantiasis, hydroceole	infected
Brugia malayi	South and East Asia	*Mansonia* and Anopheline mosquitoes	lymphatics	blood	lymphangitis, elephantiasis	infected
Brugia timori	Indonesia	Anopheline mosquitoes	lymphatics	blood	lymphangitis, elephantiasis	infected
Loa loa	Central and West Africa	*Chrysops* flies	connective tissue	blood	dermatitis	uninfected
Mansonella perstans	Africa, Central and South America	*Culicoides* midges	body cavities	blood	mostly asymptomatic	varies regionally
Mansonella streptocerca	Central and West Africa	*Culicoides* midges	skin	skin	mostly asymptomatic	unknown
Mansonella ozzardi	Central and South America	*Culicoides* midges and *Simulium* flies	subcutaneous tissues	blood	mostly asymptomatic	infected
Onchocerca volvulus	Africa (99%), Yemen, Central and South America	*Simulium* flies	subcutaneous nodules	skin	dermatitis, eye disease (river blindness)	infected

Table 3: *Wolbachia* in filarial nematodes. This information was adapted from a table provided by J. A. Vaughan, University of North Dakota.

Subfamily	Species	*Wolbachia* status	Reference
Dirofilariinae	*Dirofilaria immitis*	+	[41,46]
Dirofilariinae	*Dirofilaria repens*	+	[46]
Dirofilariinae	*Foleyella candezei*	-	[32]
Dirofilariinae	*Foleyella furcata*	-	[51]
Dirofilariinae	*Loa loa*	-	[47,49]
Onchocercinae	*Acanthocheilonema reconditum*	-	[51]
Onchocercinae	*Acanthocheilonema viteae*	-	[41,46]
Onchocercinae	*Brugia malayi*	+	[96]
Onchocercinae	*Brugia pahangi*	+	[41,46]
Onchocercinae	*Cercopithifilaria bulboidea*	-	[32]
Onchocercinae	*Cercopithifilaria crassa*	-	[32]
Onchocercinae	*Cercopithifilaria japonica*	+	[32]
Onchocercinae	*Cercopithifilaria longa*	-	[32]
Onchocercinae	*Cercopithifilaria minuta*	-	[32]
Onchocercinae	*Cercopithifilaria multicauda*	-	[32]
Onchocercinae	*Cercopithifilaria roussilhoni*	-	[32]
Onchocercinae	*Cercopithifilaria shohoi*	-	[32]
Onchocercinae	*Cercopithifilaria tumidicervicata*	-	[32]
Onchocercinae	*Dipetalonema gracile*	+	[51]
Onchocercinae	*Dipetalonema setariosum*	-	[41]
Onchocercinae	*Litomosa chiropterorum*	-	[97]
Onchocercinae	*Litomosa westi*	+	[51]
Onchocercinae	*Litomosoides brasiliensis*	+	[51]
Onchocercinae	*Litomosoides galizai*	+	[51]
Onchocercinae	*Litomosoides hamletti*	+	[51]
Onchocercinae	*Litomosoides sigmodontis*	+	[46]
Onchocercinae	*Litomosoides taylori*	+	[32]
Onchocercinae	*Litomosoides yutajensis*	-	[51]
Onchocercinae	*Loxodontofilaria caprini*	mixed	[32]
Onchocercinae	*Mansonella atelensis*	+	[32]
Onchocercinae	*Mansonella ozzardi*	+	[51]
Onchocercinae	*Mansonella perforata*	mixed	[32]
Onchocercinae	*Mansonella perstans*	mixed	[52,98]
Onchocercinae	*Monanema martini*	-	[32]
Onchocercinae	*Onchocerca armillata*	+	[45]
Onchocercinae	*Onchocerca diwettei*	mixed	[32]
Onchocercinae	*Onchocerca eberhardi*	+	[32]

Onchocercinae	*Onchocerca fasciata*	+	[99]
Onchocercinae	*Onchocerca flexuosa*	-	[50]
Onchocercinae	*Onchocerca gibsoni*	+	[46]
Onchocercinae	*Onchocerca gutturosa*	+	[46]
Onchocercinae	*Onchocerca jakutensis*	+	[50]
Onchocercinae	*Onchocerca linealis*	+	[45]
Onchocercinae	*Onchocerca ochengi*	+	[46]
Onchocercinae	*Onchocerca skjabini*	+	[32]
Onchocercinae	*Onchocerca suzukii*	mixed	[32]
Onchocercinae	*Onchocerca tarsicola*	+	[45]
Onchocercinae	*Onchocerca volvulus*	+	[40]
Onchocercinae	*Wuchereria bancrofti*	+	[46]
Oswaldofilarinae	*Piratuba scaffi*	-	[32]
Setariinae	*Setaria digitata*	-	[32]
Setariinae	*Setaria equina*	-	[48]
Setariinae	*Setaria labiatopapillosa*	-	[51]
Setariinae	*Setaria* sp.	-	[32]
Setariinae	*Setaria* sp. 2	-	[32]
Setariinae	*Setaria* sp. 3	-	[32]
Setariinae	*Setaria tundra*	-	[51]
Splendidofilariinae	*Aproctella* sp. 1	-	[32]
Splendidofilariinae	*Chandlerella quiscali*	-	[100]
Waltonellinae	*Ochoterenella royi*	-	[32]
Waltonellinae	*Ochoterenella* sp.	-	[51]
Waltonellinae	*Ochoterenella* sp. 1	-	[32]

Table 4: Comparison of doxycycline with traditional anthelmintics. Information was taken from Gyapong et al., 2005; Fox, 2006; Basanez et al. 2008; Hoerauf, 2008; Taylor et al., 2010; Osei-Atweneboana et al. [19,64,101,102,103,104].

	Doxycycline	**Ivermectin**	**Ivermectin + Albendazole**	**DEC + Albendazole**
Targeted diseases	Any species with *Wolbachia*	*O. volvulus*	LF in Africa	LF outside Africa
Duration of treatment	4 to 6 weeks of treatment	Single dose, repeated every 6-12 months in MDA	Single dose, repeated annually in MDA	Single dose, repeated annually in MDA
General anthelmintic activity	None	Effective against broad range of helminths	Effective against broad range of helminths	Effective against broad range of helminths
Microfilaricidal activity	Slow MF decline	Rapid MF decline	Fast MF decline	Fast MF decline
Effect on worm fertility	Long-term sterility (>2 years)	MF levels rise slowly after three months	Reduces MF levels for at least 1 year	Reduces MF levels for at least 1 year
Macrofilaricidal activity	78% reduction alone, 83-92% reduction with ivermectin and albendazole	Limited	Limited	DEC alone and DEC/Alb have partial macrofilaricidal activity
Effect on disease pathology	Reduction in lymphedema and hydrocele size	Little effect	Little effect	Little effect
Contraindications	Cannot be used in children or pregnant/nursing women	High burden of *L. loa* coinfection leads to severe adverse reactions	High burden of *L. loa* coinfection leads to severe adverse reactions	Cannot be used in areas where patients have *O. volvulus* infections
Side effects from medication	Upset stomach, sun sensitivity, staining of teeth in children	Generally well tolerated	Generally well tolerated	Generally well tolerated
Side effects from death of worms or MF	No adverse reactions	Itching, fever, etc. in patients with onchocerciasis. Fever, headache in patients with LF	Same as Ivermectin alone	Same as Ivermectin plus localized inflammation and pain near dead adult worms. Not safe for use in onchocerciasis.
Resistance	Unknown for *Wolbachia*; documented in related bacteria	Resistance reported in intestinal parasites of sheep and cattle; recent report of suspected resistance in *O. volvulus*	Reports of albendazole resistance in intestinal parasites of cattle and sheep	None documented

CHAPTER 2:

Comparing the Mitochondrial Genomes of *Wolbachia*-Dependent and Independent Filarial Nematode Species

Submitted to BMC Genomics on November 9, 2011.

Samantha N. McNulty[1], Andrew S. Mullin[1], Jefferson A. Vaughan[2], Vasyl V. Tkach[2], Gary J. Weil[1], Peter U. Fischer[1]

[1]Infectious Diseases Division, Department of Internal Medicine, Washington University School of Medicine, Campus Box 8051, 660 S. Euclid Avenue, St. Louis, MO 63110

[2]Department of Biology, University of North Dakota, 10 Cornell St, Grand Forks, ND 58202

PREFACE

SNM sequenced the mitochondrial genomes of *Acanthocheilonema viteae*, *Chandlerella quiscali* and *Onchocerca flexuosa* and performed assembly, annotation and bioinformatic analyses on the mitochondrial genomes described in this chapter. SNM wrote the first complete draft of this chapter. Comments from Gary J. Weil and Peter U. Fischer were incorporated into the final version presented here.

ABSTRACT

Background: Many species of filarial nematodes depend upon *Wolbachia* endobacteria to carry out their life cycles. Other filarial species are naturally *Wolbachia*-free. Careful comparisons of *Wolbachia*-dependent and independent species, including comparisons of their mitochondrial genomes, may shed light on the reason(s) for this disparity.

Results: The mitochondrial genomes of five species of filarial nematodes, namely *Acanthocheilonema viteae, Chandlerella quiscali*, *Loa loa*, *Onchocerca flexuosa*, and *Wuchereria bancrofti*, were sequenced, annotated and compared with available mitochondrial genome sequences from *Brugia malayi, Dirofilaria immitis, Onchocerca volvulus* and *Setaria digitata*. No major differences were observed in genome size, AT content or codon usage. All nine mitochondrial genomes contain 12 protein-coding genes, 22 tRNAs and 2 rRNAs. Synteny was perfectly preserved in all species except *C. quiscali,* which had a different order for 5 tRNA genes. Homology of protein coding genes among various species was consistent with reported phylogenies of non-mitochondrial genes and not with the organism's *Wolbachia* status.

Conclusions: These results have provided new mitochondrial genome sequences from filarial nematodes. The presence of *Wolbachia* endobacteria appears to have had little impact on the structure and sequence of filarial mitochondrial genomes, as gene content and organization are highly conserved among all examined species. Additional research will be needed to determine whether mitochondria from *Wolbachia*-dependent filarial species are functionally different from those of *Wolbachia*-independent species despite their sequence level similarities.

ABBREVIATIONS

ATP6, ATP synthase subunit 6; CytB, cytochrome b; Cox1, Cox2 and Cox3, cytochrome C oxidase subunits 1–3; ND1–6, and NDL4, nicotinamide adenine dinucleotide dehydrogenase subunits 1–6, and L4; rRNA, ribosomal RNA; tRNA, transfer RNA.

INTRODUCTION

Many filarial nematode species, including several important human pathogens, contain *Wolbachia* endobacteria that are necessary for worm growth, molting, fertility, and in some cases, survival [1,2,3]. Other filarial species are naturally *Wolbachia*-free [4,5,6,7,8,9]. This phenomenon may have been caused by or may have led to significant genetic differences between these two groups. Studies comparing gene content and expression in *Wolbachia*-infected and uninfected filarial nematodes may provide insight into the exact biological mechanisms underlying this disparity and into the evolutionary effects exerted by the *Wolbachia* endosymbiont.

Comparisons of *Wolbachia*-infected and uninfected species should include an examination of their mitochondrial genomes (mtDNA), as mitochondria are particularly sensitive to the evolutionary pressures exerted by *Wolbachia* infection. *Wolbachia* and mitochondria are co-transmitted through vertical passage from mother to offspring; therefore, the expansion of *Wolbachia* in a population is concurrent with the expansion of the mitochondrial haplotypes associated with infected individuals. Such *Wolbachia*-mitochondria "sweeps," characterized by unusually low degrees of variation in the mtDNA of infected populations, have been noted in many insect species [10,11,12,13,14]. A similar lack of mtDNA diversity is seen in populations of *Dirofilaria immitis* (canine heartworm) in comparison to *Wolbachia*-free, non-filarial nematodes [15]. In light of this phenomenon, we hypothesized that the mtDNAs of *Wolbachia*-dependent filaria may differ in gene content, sequence or organization as compared to those found in *Wolbachia*-free species whose ancestor(s) did not undergo a *Wolbachia*-induced genetic bottleneck or evolve in the presence of an endobacterial partner.

The mtDNA sequences of four species of filarial nematodes have been reported to date. These include *Onchocerca volvulus* [16], *D. immitis* [17], *Brugia malayi* [18], and *Setaria digitata* [19]. *D. immitis, O. volvulus* and *B. malayi* are *Wolbachia*-dependent. At the time its mtDNA sequence was generated, the *Wolbachia* status of *S. digitata* was unknown but it has since been characterized as *Wolbachia*-free [9]. This manuscript reports the generation of mtDNA sequences for five additional species: *Wuchereria bancrofti*, *Acanthocheilonema viteae, Loa loa, Onchocerca flexuosa* and *Chandlerella quiscali*. *W. bancrofti* is a *Wolbachia*-dependent human pathogen that infects some 115 million people throughout the tropics [20]. *A. viteae*, *L. loa* and *O. flexuosa* are *Wolbachia*-free, but they belong to subfamilies (i.e. the Onchocercinae and the Dirofilariinae) dominated by *Wolbachia*-dependent species [6,9]. In fact, *Wolbachia*-like DNA sequences in the nuclear genomes of *A. viteae* and *O. flexuosa* indicate that they were infected in the distant past [21]. *C. quiscali* is also *Wolbachia*-free [22], and like *S. digitata*, comes from a subfamily (Splendidofilariinae) that may be free of *Wolbachia* infection [9].

The purpose of the reported study was to compare mtDNA from *Wolbachia*-dependent and independent filarial species from *Wolbachia*-associated (i.e. Onchocercinae and Dirofilariinae) and unassociated (i.e. Splendidofilariinae and Setariinae) subfamilies. Our analyses revealed no major sequence-level discrepancies that could be attributed to *Wolbachia* status. Future studies will be required to discover subtler affects of *Wolbachia* on the sequence or function of filarial nematode mitochondria.

RESULTS

Gene Content and Organization

The mitochondrial genomes (mtDNA) of five species of filarial nematodes were sequenced, annotated and deposited in Genbank under the accession numbers HQ186249 for *A. viteae*, HM773029 for *C. quiscali*, HQ186250 for *L. loa*, HQ214004 for *O. flexuosa* and HQ184469 for *W. bancrofti*. The general properties of these mtDNAs are compared to those of *B. malayi*, *D. immitis, O. volvulus* and *S. digitata* in **Table 1**. The newly sequenced mtDNAs are similar in size and AT content to those of other filarial species. So far, filarial mtDNAs range in size from 13,474 bp in *O. volvulus* to 13,839 in *S. digitata* and range in AT content from 73.7% in *O. volvulus* to 77.7% in *C. quiscali* [16,19].

All nine filarial mtDNAs encode the same 12 proteins, 22 tRNAs and 2 rRNAs (**Fig. 1**, **Table 1**) with very short intergenic sequences. All genes are encoded in the same direction, a characteristic shared by most nematode mtDNAs. Synteny is perfectly preserved in all examined species with the exception of *C. quiscali* (**Fig. 1**). In eight of the nine examined species, five tRNA genes ($tRNA^{Ala}$, $tRNA^{Leu2}$, $tRNA^{Asn}$, $tRNA^{Met}$ and $tRNA^{Lys}$) are clustered together between the AT-rich region and the NDL4 gene. However, in *C. quiscali*, the $tRNA^{Met}$ gene is positioned between the Cox1 gene and AT-rich region separate from the main tRNA cluster, and the order of the other of the other four tRNA genes are scrambled relative to other species.

Protein Coding Genes

The 12 protein coding genes were identified in *A. viteae*, *C. quiscali*, *L. loa*, *O. flexuosa* and *W. bancrofti* based on homology with orthologs from the other species. None of these genes contain premature stop codons or frameshift mutations. RT-PCR reactions indicate that the predicted protein coding genes were expressed at the RNA level in all examined species (**Fig. 2**).

Filarial mtDNAs are extremely thymine (T)-rich (**Table 1**). In light of this, it is not surprising that filarial mitochondria show a bias towards T-rich codons (**Table S1**). The most frequently used codon in all species is TTT, which encodes phenylalanine and serves as an alternative start codon in certain instances (**Table S1**, **Table 2**). The start and stop codons used by each species are listed in **Table 2**. Novel start codons include TGT for ND6 in *W. bancrofti*, TCT for CytB in *A. viteae*, and CCT for ND3 in *O. flexuosa*. Termination codons include TAG, TAA, and the incomplete stop codon T. The T codon is converted to TAA upon addition 3' poly(A) tail.

Neighbor joining phylogenetic trees were constructed to examine homology between the protein coding genes of the various species (**Fig. 3**). Trees were constructed from concatenated protein coding sequences (**Fig. 3a**) as well as single protein coding sequences, though only those based on Cox1 (**Fig. 3b**) and Cox2 (**Fig. 3c**) are reported. Although bootstrap values varied, all trees had similar clustering patterns. Overall, clustering indicated that the most closely related species pairs were *O. volvulus* and *O. flexuosa* and *B. malayi* and *W. bancrofti*. In all trees, *Chandlerella quiscali* (the sole member of the Splendidofilariinae in the analysis) was basal to the remaining taxa. In no case did the clustering appear to be reflective of *Wolbachia* infection status. *Wolbachia*-

dependent and *Wolbachia*-independent filarial species do not form monophyletic lineages in any of the gene trees or the combined analysis.

Ribosomal and Transfer RNA Genes

Like the protein coding genes, the rRNA genes of the five species sequenced in this study were identified based on homology with known orthologs. In all species, the 12s rRNA gene is positioned between NDL4 and ND1 while the 16s rRNA gene is positioned between Cox2 and ND3 (**Fig. 1**). The exact boundaries of these genes have yet to be mapped in any filarial species.

Filarial tRNAs were identified based on predicted secondary structure and by homology to tRNAs from previously sequenced filarial mtDNAs. In the previously sequenced species, 20 of the 22 mitochondrial tRNAs share a common secondary structure in which the TΨC arm and variable loop are exchanged for a TV-replacement loop [16,17]. Conversely, the two $tRNA^{Ser}$ genes contain a DHU replacement loop in exchange for the typical D arm (**Fig. 4**) [16,17]. The predicted mitochondrial tRNA structures of *A. viteae* followed this trend exactly, as did most of the tRNAs from the other examined species. However, our predictions indicate that $tRNA^{Ser1}$ and $tRNA^{Asn}$ in *C. quiscali*, $tRNA^{Lys}$ and $tRNA^{Pro}$ in *L. loa*, $tRNA^{Trp}$ in *O. flexuosa* and $tRNA^{Pro}$ in *W. bancrofti* may contain both the TΨC and D loops (**Fig. 4**). The same anticodons are used in all species with two exceptions. $tRNA^{Pro}$ uses the anticodon AGG in *O. volvulus*, *D. immitis*, *S. digitata* and *O. flexuosa* while the anticodon TGG is used in other species, and $tRNA^{Leu1}$ uses the anticodon TAA in *A. viteae* while the anticodon TAG is used in other species.

AT rich region

The control, or AT rich, region represents the largest non-coding region in filarial mtDNAs, which are otherwise densely packed with tightly spaced or slightly overlapping protein coding, tRNA and rRNA genes. The AT rich regions of the 9 sequenced filarial mtDNAs range in size from 256 bp in *W. bancrofti* to 506 bp in *S. digitata* (**Table 1**). In most species, this region is located between the genes encoding Cox3 and $tRNA^{Ala}$. The unusual arrangement of tRNA genes in *C. quiscali* places its proposed 308 bp AT rich region between the $tRNA^{Met}$ and $tRNA^{Leu2}$ genes, leaving an additional 109 bp non-coding region between the Cox3 and $tRNA^{Met}$ genes (**Fig. 1**) The impact of this secondary non-coding region is unknown.

DISCUSSION

Mitochondria serve as the powerhouse of the eukaryotic cell. Because of their vital function, their small genome size, and the presence of well-established phylogenetic markers, mtDNA sequences are often among the first that are obtained in genomic studies of new organisms. To date, 1511 metazoan mitochondrial genomes have been sequenced and are available in public databases (see http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/mztax_short.html). Despite their obvious socioeconomic and medical importance, only four of these mitochondrial genomes are from filarial nematodes [16,17,18,19]. Our study more than doubles the number of sequenced filarial mitochondrial genomes available in the public databases.

We used the newly available sequences to compare the mtDNA of *Wolbachia*-dependent and independent filarial nematodes. Initially, we hypothesized that there might be obvious differences in the mtDNA sequences of *Wolbachia*-dependent and independent species due to the evolutionary pressure exerted by the endosymbiont [10,11,12,13,14,15]. However, it appears that the mtDNAs of filarial nematodes are remarkably similar. No major differences in genome length, AT content or codon usage were detected. All nine species contain the standard 12 protein coding genes, 22 tRNAs and 2 rRNAs with synteny perfectly preserved in all species except *C. quiscali*. The rearrangement of five tRNA's in *C. quiscali* probably reflects its evolutionary distance from the other species included in this study rather than its *Wolbachia*-free status (see **Fig. 3**), as such rearrangements were not detected in other *Wolbachia*-free species. Mitochondrial genomes of other members of the Splendidofilariinae need to be sequenced in order to find out if the observed difference in the gene order is unique to

this species, or it is typical of the whole subfamily. Differences in tRNA structure or anticodon usage are also minor and do not correlate to *Wolbachia* status. In light of these findings, it seems that *Wolbachia* has had little effect on the content, arrangement and sequence of filarial mtDNA.

If *Wolbachia* infection had led to the fixation of certain mitochondrial types in an ancestral population, one might expect to see higher degrees of sequence homology between infected daughter species relative to uninfected species, particularly species like *C. quiscali* and *S. digitata* that may have never come into contact with the bacteria [9]. However, none of the mitochondrial protein coding genes cluster based on *Wolbachia* dependence, regardless of the species' belonging to a *Wolbachia*-associated or unassociated subfamily (**Fig. 3**). In fact, the Neighbor Joining trees generated in this study are remarkably similar to those generated from non-mitochondrial markers like the 5s intergenic spacer [23,24]. Our assessment of homology was done at the level of complete protein coding genes. It may be more informative to assess smaller loci, as even single base changes are known to have profound effects on mitochondrial function [25]. Of course, substantial differences may yet exist in nuclear genes related to mitochondrial function, but these will not be discovered until more filarial genomes are sequenced.

Although no consistent sequence-level differences were observed between mitochondria from *Wolbachia*-dependent and independent filarial species, this does not exclude the possibility that differences exist in mitochondrial function or efficiency. For example, Strubling *et al*. have reported that tetracycline-mediated clearance of the *Wolbachia* endosymbiont led to increased expression of 10 of 12 mitochondrial protein coding genes in the *Wolbachia*-dependent filarial species *Litomosoides sigmodontis*,

while the same treatment did not affect the expression of these genes in *A. viteae* [26]. If the expression of mitochondrial protein coding genes can be altered by the presence or absence of *Wolbachia*, it is possible that these genes could be expressed at different levels in *Wolbachia*-dependent versus independent species, leading to different levels of mitochondrial efficiency. Of course, RNA and protein expression levels may correlate to the number of mitochondrial genomes present in each mitochondria or the number of mitochondria in a given cell or organism. These factors may also vary between the two groups.

Mitochondria and *Wolbachia* share many common features including rickettsial ancestry [27,28], reduced genome content [27,29,30], transfer of DNA to the nucleus [21,31,32], and maternal transmission [3]. Over time, the presence of *Wolbachia* could have affected the function of mitochondria in *Wolbachia*-dependent species without having major effects on the mtDNA sequence. Additional research will be required to test this hypothesis.

MATERIALS AND METHODS

Parasite materials

Adult *B. malayi* and *A. viteae* were obtained from experimentally infected Mongolian jirds as previously described [33,34]. Adult *D. immitis* were obtained from the Filariasis Research Reagent Resource Center (Athens, GA). Adult *O. flexuosa* were isolated from subcutaneous nodules dissected from red deer (*Cervus elaphus*) in Schleswig-Holstein, Germany [8]. Adult *O. volvulus*, and microfilariae of *W. bancrofti* and *L. loa* were available from prior studies in Uganda, Papua New Guinea and Cameroon, respectively [5,35,36]. Adult *Chandlerella quiscali* were obtained from common grackles *Quiscalus quiscula* trapped in North Dakota, USA.

Nucleic acid isolation and cDNA synthesis

DNA for sequencing was isolated from adult worms and microfilariae using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA), ethanol precipitated to concentrate and stored in 1x TE buffer. RNA was isolated as previously described [21]. Briefly, worms were homogenized by bead-beating in TRIzol (Invitrogen, Carlsbad, CA) and RNA was isolated by organic extraction with 1-bromo-3-chloropropane followed by column purification using the RNeasy Mini Kit (Qiagen) including an on-column DNase digest. A second DNase treatment was performed with the TURBO DNA-*free* Kit (Applied Biosystems, Austin, TX). cDNA was synthesized from total RNA using qScript cDNA SuperMix according to manufacturer's suggested protocol (Quant Biosciences, Gaithersburg, MD) and purified with the Qiaquick PCR Purification Kit (Qiagen).

PCR reactions and sequencing

Primers used to amplify mtDNA are reported in **Table S2**. PCR products were cloned using the TOPO-TA Cloning Kit for Sequencing or the TOPO-XL PCR Cloning Kit (Invitrogen) depending on size, and sequenced by primer walking.

Primers used to detect expression of mitochondrial genes by RT-PCR are reported in **Table S2**. Each of these reactions included a DNA positive control, a cDNA test sample and total RNA and water-only negative controls.

Assembly and annotation of the mitochondrial genomes

Sequenced contigs were assembled using Contig Express and analyzed using Vector NTI version 10.3.1 (Invitrogen). Sequences were verified by comparison with publically available sequence data from the Genbank sequence read archive for *L. loa* (accession number SRP000756) and *W. bancrofti* (accession number SRP000772).

Protein coding genes (including initiation and termination codons) and rRNAs were determined based on their homology to sequences reported from the mitochondrial genomes of *B. malayi, D. immitis, O. volvulus* and *S. digitata* [16,17,18,19].

In most instances, tRNA sequences were predicted using the Arwen program (available at http://130.235.46.10/ARWEN/) [37] and verified by homology to known filarial tRNA sequences. Any computationally predicted tRNAs that fell within other documented structures (i.e. protein coding genes or rRNAs) were disregarded. $tRNA^{ala}$ and $tRNA^{leu2}$ in *A. viteae*, $tRNA^{leu2}$ and $tRNA^{gly}$ in *O. flexuosa*, and $tRNA^{ala}$ in *L. loa* were identified solely based on homology to known orthologs and the presence of the expected anticodon.

AT percentages and codon usage were calculated using the DNA Stats and codon usage features available from the Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/dna_stats.html). Diagrams of complete mitochondrial genomes were constructed using DNA plotter (http://www.sanger.ac.uk/resources/software/dnaplotter/) [38].

Phylogenetic trees

Phylogenetic analyses were conducted in MEGA4 [39] using the Neighbor-Joining method [40] with sequences from *Caenorhabditis elegans* serving as the outgroup. Evolutionary distances were computed using the Maximum Composite Likelihood method [41]. The optimal trees with the shortest sum branch length were obtained for each dataset. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (2000 replicates) is shown next to the branches [42]. For the concatenated sequence tree, the 12 protein coding genes were assembled in the order in which they are present in filarial nematodes. For single gene trees, the full coding sequence was used for each species.

ACKNOWLEDGEMENTS

We would like to thank Richard Lucius (Humboldt University, Berlin), the Filariasis Research Reagent Resource Center (Atlanta, Georgia), Samuel Wanji (University of Buea, Cameroon), Norbert Brattig (Bernard Nocht Institute for Tropical Medicine, Hamburg) and Yuefang Huang (Washington University School of Medicine, St. Louis) for providing parasite material.

This research was funded by The Barnes Jewish Hospital Foundation. SNM was supported by NIH grant T32-AI007172 and JAV was supported by NIH grant R03-AI092306.

REFERENCES

1. Hoerauf A, Mand S, Fischer K, Kruppa T, Marfo-Debrekyei Y, et al. (2003) Doxycycline as a novel strategy against bancroftian filariasis-depletion of *Wolbachia* endosymbionts from *Wuchereria bancrofti* and stop of microfilaria production. Med Microbiol Immunol 192: 211-216.
2. Hoerauf A, Nissen-Pahle K, Schmetz C, Henkle-Duhrsen K, Blaxter ML, et al. (1999) Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility. J Clin Invest 103: 11-18.
3. Taylor MJ, Hoerauf A (1999) *Wolbachia* bacteria of filarial nematodes. Parasitol Today 15: 437-442.
4. Bandi C, Anderson TJ, Genchi C, Blaxter ML (1998) Phylogeny of *Wolbachia* in filarial nematodes. Proc Biol Sci 265: 2407-2413.
5. Buttner DW, Wanji S, Bazzocchi C, Bain O, Fischer P (2003) Obligatory symbiotic *Wolbachia* endobacteria are absent from *Loa loa*. Filaria J 2: 10.
6. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. Int J Parasitol 34: 191-203.
7. McGarry HF, Pfarr K, Egerton G, Hoerauf A, Akue JP, et al. (2003) Evidence against *Wolbachia* symbiosis in *Loa loa*. Filaria J 2: 9.
8. Plenge-Bonig A, Kromer M, Buttner DW (1995) Light and electron microscopy studies on *Onchocerca jakutensis* and *O. flexuosa* of red deer show different host-parasite interactions. Parasitol Res 81: 66-73.
9. Ferri E, Bain O, Barbuto M, Martin C, Lo N, et al. (2011) New Insights into the Evolution of *Wolbachia* Infections in Filarial Nematodes Inferred from a Large Range of Screened Species. PLoS One 6: e20843.
10. Delgado AM, Cook JM (2009) Effects of a sex-ratio distorting endosymbiont on mtDNA variation in a global insect pest. BMC Evol Biol 9: 49.
11. Hurst GD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc Biol Sci 272: 1525-1534.
12. Raychoudhury R, Grillenberger BK, Gadau J, Bijlsma R, van de Zande L, et al. (2010) Phylogeography of *Nasonia vitripennis* (Hymenoptera) indicates a mitochondrial-*Wolbachia* sweep in North America. Heredity 104: 318-326.
13. Turelli M, Hoffmann AA (1991) Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature 353: 440-442.
14. Turelli M, Hoffmann AA, McKechnie SW (1992) Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. Genetics 132: 713-723.
15. Belanger DH, Perkins SL (2010) *Wolbachia* infection and mitochondrial diversity in the canine heartworm (*Dirofilaria immitis*). Mitochondrial DNA 21: 227-233.
16. Keddie EM, Higazi T, Unnasch TR (1998) The mitochondrial genome of *Onchocerca volvulus*: sequence, structure and phylogenetic analysis. Mol Biochem Parasitol 95: 111-127.

17. Hu M, Gasser RB, Abs El-Osta YG, Chilton NB (2003) Structure and organization of the mitochondrial genome of the canine heartworm, *Dirofilaria immitis*. Parasitology 127: 37-51.
18. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al. (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317: 1756-1760.
19. Yatawara L, Wickramasinghe S, Rajapakse RP, Agatsuma T (2010) The complete mitochondrial genome of *Setaria digitata* (Nematoda: Filarioidea): Mitochondrial gene content, arrangement and composition compared with other nematodes. Mol Biochem Parasitol 173: 32-38.
20. Michael E, Bundy DA (1997) Global mapping of lymphatic filariasis. Parasitol Today 13: 472-476.
21. McNulty SN, Foster JM, Mitreva M, Dunning Hotopp JC, Martin J, et al. (2010) Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. PLoS One 5: e11029.
22. McNulty SN, Fischer K, Mehus JO, Vaughan JA, Tkach VV, et al. (2011) Absence of *Wolbachia* Endobacteria in *Chandlerella quiscali*: An Avian Filarial Parasite. J Parasitol.
23. Michalski ML, Bain O, Fischer K, Fischer PU, Kumar S, et al. (2010) Identification and phylogenetic analysis of *Dirofilaria ursi* (Nematoda: Filarioidea) from Wisconsin black bears (*Ursus americanus*) and its *Wolbachia* endosymbiont. J Parasitol 96: 412-419.
24. Xie H, Bain O, Williams SA (1994) Molecular phylogenetic studies on filarial parasites based on 5S ribosomal spacer sequences. Parasite 1: 141-151.
25. Wong LJ (2010) Molecular genetics of mitochondrial disorders. Dev Disabil Res Rev 16: 154-162.
26. Strubing U, Lucius R, Hoerauf A, Pfarr KM (2010) Mitochondrial genes for heme-dependent respiratory chain complexes are up-regulated after depletion of *Wolbachia* from filarial nematodes. Int J Parasitol 40: 1193-1202.
27. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, et al. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396: 133-140.
28. Taylor MJ, Bandi C, Hoerauf A (2005) *Wolbachia* bacterial endosymbionts of filarial nematodes. Adv Parasitol 60: 245-284.
29. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, et al. (2005) The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. PLoS Biol 3: e121.
30. Karlberg O, Canback B, Kurland CG, Andersson SG (2000) The dual origin of the yeast mitochondrial proteome. Yeast 17: 170-187.
31. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753-1756.
32. Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol 60: 115-138.
33. Ash LR, Riley JM (1970) Development of subperiodic *Brugia malayi* in the jird, *Meriones unguiculatus*, with notes on infections in other rodents. J Parasitol 56: 969-973.

34. Lucius R, Textor G (1995) *Acanthocheilonema viteae*: rational design of the life cycle to increase production of parasite material using less experimental animals. Appl Parasitol 36: 22-33.
35. Fischer P, Kipp W, Bamuhiga J, Binta-Kahwa J, Kiefer A, et al. (1993) Parasitological and clinical characterization of *Simulium neavei*-transmitted onchocerciasis in western Uganda. Trop Med Parasitol 44: 311-321.
36. Weil GJ, Kastens W, Susapu M, Laney SJ, Williams SA, et al. (2008) The impact of repeated rounds of mass drug administration with diethylcarbamazine plus albendazole on bancroftian filariasis in Papua New Guinea. PLoS Negl Trop Dis 2: e344.
37. Laslett D, Canback B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24: 172-175.
38. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25: 119-120.
39. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.
40. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.
41. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101: 11030-11035.
42. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791.

FIGURES

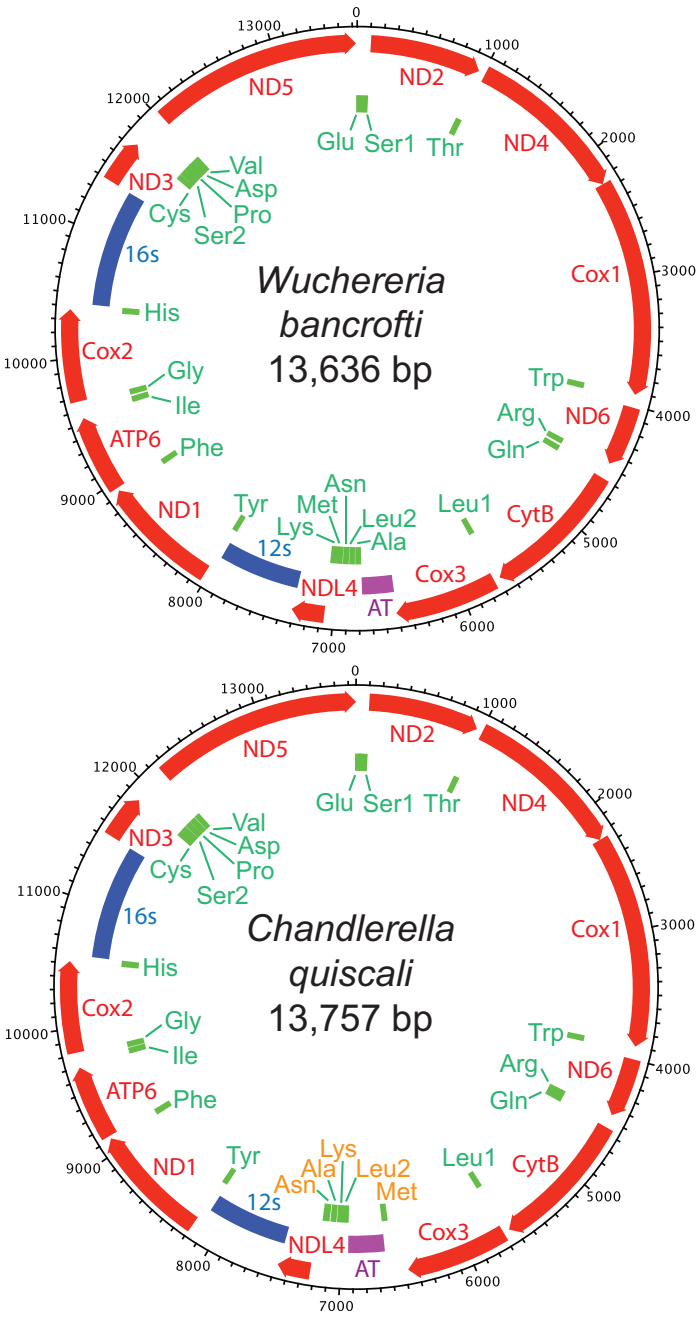


Figure 1: Comparative diagrams of the mitochondrial genomes of (**A**) *W. bancrofti* and (**B**) *C. quiscali*. Protein coding genes are shown in red with arrowheads indicating directionality. rRNA and tRNA genes are shown in blue and green, respectively, and the AT-rich region is shown in purple. The diagram of the *W. bancrofti* mitochondrial genome is representative of most filarial mitochondria, as synteny is preserved in all species except *C. quiscali*. The five tRNA genes rearranged in *C. quiscali* are highlighted in orange.

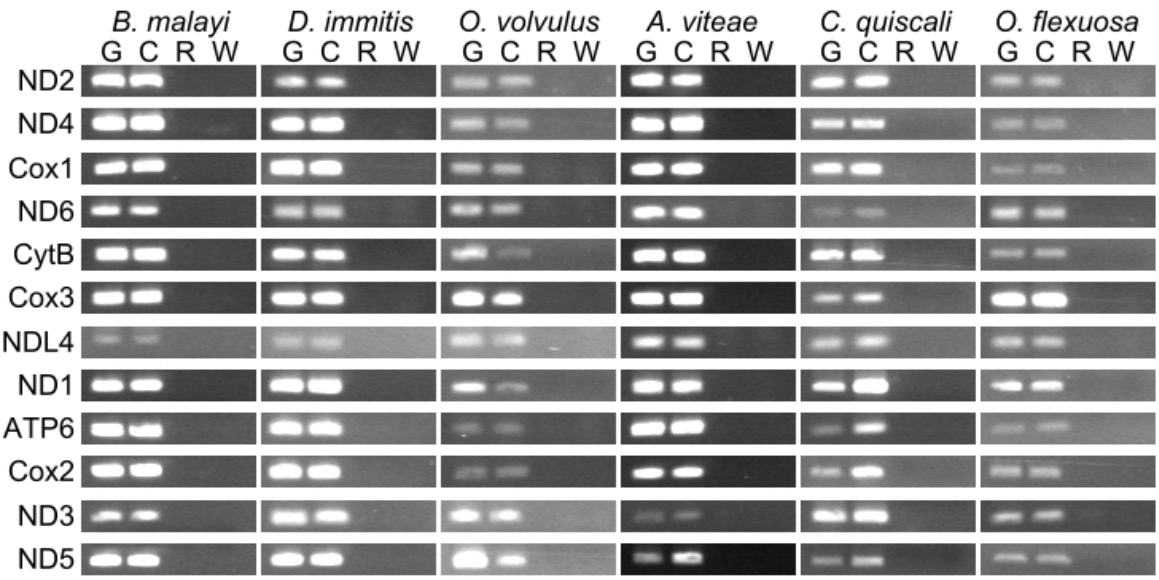


Figure 2: Expression of mitochondrial protein coding genes in six filarial nematode species. The following templates were used for each RT-PCR reaction: genomic DNA (G), cDNA (C), total RNA (R) and water (W).

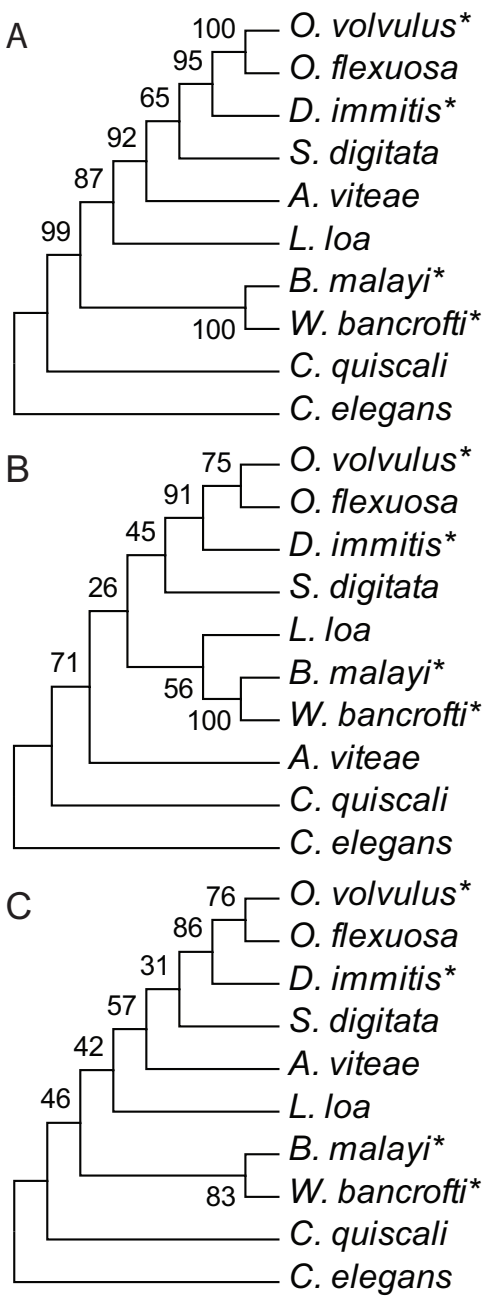
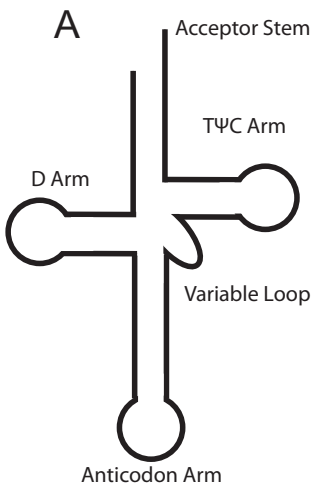


Figure 3: Phylogenetic affinities of filarial nematodes based on mitochondrial protein coding genes. Neighbor Joining trees were constructed for (**A**) concatenated protein coding genes, (**B**) Cox1, and (**C**) Cox2. Branch support was calculated using the bootstrap method (2000 replicates) with percentages displayed at nodes. Asterisk indicates that the species is *Wolbachia*-dependent.




```
B
      t
    g+t
    t-a
    a-t
    g-c
    g+t
    t.t
    t.t
   t   t
 aa    a     t
t  ttgg      t
t  !!!+      a
 t aact     g
  a    a   a
     t-a
     t-a
     a-t
     g-c
     t-a
    t   a
    t   a
     tca
```

```
C
      t
    g+t
    a-t
    g-c
    t-a
    t-a
    a-t
    a-t     t
       ttcgt
   t   !+!++ t
  t    aggtg t
       t    t
  t   t
  t   t
   t.t
   a-t
   a-t
   g-c
   t-a
  t   a
  t   a
   tga
```

```
D
      t
    g+t
    a-t
    a-t
    t.t
    t-a
    t+g
    t-a    ct
   t   aagt  t
 aa    a    !!!!  a
t  tatg     ttca  a
t  +!!+     g   ta
t  gtat    g
 tt    a   g
     a-t
     c-g
     t-a
     g+t
     a-t
    t   a
    t   g
     gtt
```

```
E
      t
    a-t
    t.t
    t-a
    a-t
    a-t
    a-t
    a-t     gc
   g   tatct  t
 attat    a    +!+!!  c
t     aatt     gtgga  c
t     !!!!    a     at
a     ttaa     g
atagat    t     t
     t-att
     t-a
     t-a
     a-t
     g-c
     t-a
    c   a
    t   a
     tct
```

```
F
      a
    a-t
    a-t
    a-t
    a-t
    g+t
    g.a
    a-t
    t+g   t
   t   gat t
 aa    a    +!! c
a  tttt     tta t
t  !+!!     t  t
 t agaa    g
  t    t   g
     t-a
     t-a
     t-a
     t-a
     g-c
     a-t
    c   a
    t   a
     ctt
```

```
G
      t
    t-a
    a-t
    c-g
    t-a
    g+t
    a-t
    a-t
    t-a    tt
   t   gaga  g
 aa    a    +!+!  a
g  tttg     tttt  c
t   !!!+    t    tt
t   aaat     t
 tta    t     a
     a-tgg
     a-t
     a-t
     t-a
     g-c
    t   g
    t   g
     tgg
```

```
H
      g
    t-a
    t-a
    t+g
    a-t
    g+t
    g+t
    t-a   t
   t   tct a
 t    a    !!!  t
  ttga     aga  t
t !!!!        tt
  aact
 a    a
     t-a
     t+g
     a-t
     g-c
     t-a
    t   a
    t   a
     tca
```

```
I
      t
    t-a
    g+t
    c-g
    t-a
    g+t
    a-t
    a-t
    t-a     tt
   t   gaaaa  t
 aa    a    +!!!!  t
t  tttg     ttttt  g
t   !!!+    t     ca
t   aaat     t
 tta    a     a
     a-tag
     a-t
     a-t
     t.t
     g-c
    t   g
    t   g
     tgg
```

Figure 4: Mitochondrial tRNA structures. The features of a typical tRNA include the acceptor stem, D arm, TΨC arm, variable loop and anticodon arm (**A**). Most filarial nematode mitochondrial tRNAs have a D arm, but a TV replacement loop is found in place of the TΨC arm and variable loop. $tRNA^{Trp}$ of *A. viteae* is shown as an example of the standard structure (**B**). $tRNA^{Ser1}$ and $tRNA^{Ser2}$ are notable exceptions to the typical structure. As seen in $tRNA^{Ser2}$ from *A. viteae*, these two contain a DHU replacement loop in lieu of a D arm (**C**). Other exceptions may include (**D**) $tRNA^{Asn}$ and (**E**) $tRNA^{Ser1}$ from *C. quiscali*, (**F**) $tRNA^{Lys}$ and (**G**) $tRNA^{Pro}$ from *L. loa*, (**H**) $tRNA^{Trp}$ from *O. flexuosa*, and (**I**) $tRNA^{Pro}$ from *W. bancrofti*, as these structures are predicted to include both the D and TΨC arms.

TABLES

Table 1: General characteristics of filarial nematode mitochondrial genomes. Information was taken from previous studies for *B. malayi* [18], *D. immitis* [17], *O. volvulus* [16], and *S. digitata* [19]. Information regarding the *Wolbachia* status of species from various subfamilies is reported in [9].

	Wolbachia-dependent species				*Wolbachia*-independent species, *Wolbachia*-associated subfamily			*Wolbachia*-independent species, *Wolbachia*-unassociated subfamily	
Species	*B. malayi*	*D. immitis*	*O. vovlulus*	*W. bancrofti*	*A. viteae*	*L. loa*	*O. flexuosa*	*C. quiscali*	*S. digitata*
Ascession Number	NC_004298	NC_005305	NC_001861	HQ184469	HQ186249	HQ186250	HQ214004	HM773029	NC_014282
Length (bp)	13,657	13,814	13,474	13,636	13,724	13,590	13,672	13,757	13,839
# proteins	12	12	12	12	12	12	12	12	12
% protein coding	76.0%	75.2%	77.2%	76.3%	75.7%	76.4%	76.1%	75.4%	75.2%
#rRNAs	2	2	2	2	2	2	2	2	2
#tRNAs	22	22	22	22	22	22	22	22	22
Length of AT rich region (bp)	283	362	312	256	421	288	284	308	506
A%	21.6%	19.3%	19.3%	20.5%	19.6%	20.8%	20.3%	23.0%	19.4%
T%	53.9%	54.9%	54.0%	54.1%	54.0%	54.8%	53.9%	54.7%	55.7%
G%	16.8%	19.3%	19.8%	18.0%	19.3%	17.7%	18.6%	15.9%	18.2%
C%	7.7%	6.5%	6.9%	7.4%	7.2%	6.7%	7.2%	6.4%	6.7%
AT%	75.5%	74.2%	73.3%	74.6%	73.5%	75.6%	74.2%	77.7%	75.1%

Table 2: Start and stop codons used in mitochondrial protein coding genes of filarial nematodes. Information was taken from previous studies for *B. malayi* [18], *D. immitis* [17], *O. volvulus* [16], and *S. digitata* [19]. Information regarding the *Wolbachia* status of species from various subfamilies is reported in [9].

	Wolbachia-dependent species				*Wolbachia*-independent species, *Wolbachia*-associated subfamily			*Wolbachia*-independent species, *Wolbachia*-unassociated subfamily	
Protein	*B. malayi*	*D. immitis*	*O. volvulus*	*W. bancrofti*	*A. viteae*	*L. loa*	*O. flexuosa*	*C. quiscali*	*S. digitata*
ND2	TTA/T	ATT/TAG	ATT/TAG	TTA/T	TTT/TAG	ATT/TAA	ATT/TAG	ATT/TAG	TTT/TAG
ND4	TTG/TAA	TTG/TAG	TTG/TAA	TTG/TAA	ATG/TAA	TTG/TAG	TTG/TAG	TTG/TAA	ATG/TAA
COX1	ATT/TAG	ATT/TAG	ATT/TAG	ATT/TAA	GTT/T	GTT/T	ATT/TAA	TTG/TAA	ATT/TAG
ND6	TAT/TAA	TAT/TAG	ATT/TAG	TGT/TAA	TAT/TAG	TAT/TAG	ATT/TAA	TTG/TAG	TTG/TAA
CYTB	ATT/T	GTT/T	ATT/TAA	ATT/T	TCT/T	ATT/T	ATT/TAA	ATT/T	GTT/T
COX3	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAG	ATT/TAA	ATA/T
NDL4	GTA/TAA	GTA/TAA	TTG/TAA	GTA/TAA	GTA/T	GTA/T	TTA/T	GTT/TAA	TTG/T
ND1	TTG/T	TTG/T	TTG/T	TTG/TAA	TTG/T	TTG/T	TTG/T	TTG/T	TTG/TAA
ATP6	ATT/TAG	TTG/TAA	ATT/TAG	ATT/TAG	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAG	TTT/TAG
COX2	ATT/TAA	ATT/T	ATT/TA	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAA	ATT/TAG
ND3	CTT/TAG	CTT/T	CTT/TAG	CTT/T	CTT/T	CTT/T	CCT/T	CTT/TAG	TTT/T
ND5	TTT/TAG	TTG/TAG	TTG/TAG	TTT/TAG	TTT/TAG	TTT/TAG	TTA/TAA	TTT/T	TTT/TAG

See the supplementary data CD for the following tables:

Table S1: Codon usage in the mitochondrial genomes of filarial nematodes.

Table S2: Primer sequences.

CHAPTER 3:

Endosymbiont DNA in Endobacteria-Free Filarial Nematodes Indicates Ancient Horizontal Genetic Transfer

PLoS One. 2010 Jun 9;5(6):e11029. doi: 10.1371/journal.pone.0011029.

Samantha N. McNulty[a], Jeremy M. Foster[b], Makedonka Mitreva[c], Julie C. Dunning Hotopp[d], John Martin[c], Kerstin Fischer[a], Bo Wu[b], Paul J. Davis[b], Sanjay Kumar[b], Norbert W. Brattig[e], Barton E. Slatko[b], Gary J. Weil[a], Peter U. Fischer[a]

[a] Infectious Diseases Division, Department of Internal Medicine, Washington University School of Medicine, Campus Box 8051, 660 S. Euclid Avenue, St. Louis, MO 63110, USA

[b] New England Biolabs, 240 County Road, Ipswich, MA 01938, USA

[c] The Genome Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63110, USA

[d] Institute for Genome Sciences, Department of Microbiology and Immunology, University of Maryland Baltimore, 801 W. Baltimore Street, Baltimore, MD, 21201

[e] Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Str. 74, 20359 Hamburg, Germany

PREFACE

SNM assisted in the collection of *O. flexuosa* material, isolated *O. flexuosa* genomic DNA for construction of paired-end genomic libraries, mapped the regions of *O. flexuosa* and *A. viteae* contigs with homology to *Wolbachia* sequences, confirmed all *Wolbachia*-like sequences by PCR amplification and re-sequencing, assigned *Wolbachia*-like sequences to COG functional categories and performed qRT-PCR reactions to assess the expression of *Wolbachia*-like sequences. SNM wrote the first complete draft of this manuscript. Comments from co-authors and reviewers were incorporated into the published version.

ABSTRACT

Background: *Wolbachia* are among the most abundant symbiotic microbes on earth; they are present in about 66% of all insect species, some spiders, mites and crustaceans, and most filarial nematode species. Infected filarial nematodes, including many pathogens of medical and veterinary importance, depend on *Wolbachia* for proper development and survival. The mechanisms behind this interdependence are not understood. Interestingly, a minority of filarial species examined to date are naturally *Wolbachia*-free.

Methodology/Principle Findings: We used 454 pyrosequencing to survey the genomes of two distantly related *Wolbachia*-free filarial species, *Acanthocheilonema viteae* and *Onchocerca flexuosa*. This screen identified 49 *Wolbachia*-like DNA sequences in *A. viteae* and 114 in *O. flexuosa*. qRT-PCR reactions detected expression of 30 *Wolbachia*-like sequences in *A. viteae* and 56 in *O. flexuosa*. Approximately half of these appear to be transcribed from pseudogenes. *In situ* hybridization showed that two of these pseudogene transcripts were specifically expressed in developing embryos and testes of both species.

Conclusions/Significance: These results strongly suggest that the last common ancestor of extant filarial nematodes was infected with *Wolbachia* and that this former endosymbiont contributed to their genome evolution. Horizontally transferred *Wolbachia* DNA may explain the ability of some filarial species to live and reproduce without the endosymbiont while other species cannot.

INTRODUCTION

Several important evolutionary milestones, such as the emergence of eukaryotes and the development of intracellular organelles (e.g. mitochondria and plastids), have involved disparate species uniting to form composite organisms. Filarial nematodes and their *Wolbachia* endobacteria are an interesting example of such a composite [1]. Although antibiotics can be used to cure arthropods of their *Wolbachia* infection, similar treatments lead to infertility, improper development and sometimes death of *Wolbachia*-dependent filarial nematodes [2,3]. Likewise, attempts to maintain filarial *Wolbachia* in culture have failed (B.E. Slatko, P.U. Fischer, R.U. Rao, pers. comm.). This interdependence may be a consequence of reductive evolution in both partners, as several biosynthetic pathways (e.g. synthesis of heme, riboflavin, nucleotides, etc.) seem to require both genomes for complete functionality [4].

Wolbachia are vertically transmitted through infected oocytes [5]. Their presence in the germline allows for heritable DNA transfer from the bacteria to the metazoan host. Horizontal genetic transfers (HGTs) have been reported in several *Wolbachia*-infected filarial and arthropod species [6,7,8,9], and evidence for transcription of transferred sequences has been reported in *Drosophila ananassae* that have been cured of *Wolbachia* by antibiotic treatment [7]. Similar expression studies in filarial nematodes are difficult because the infection cannot be cleared without damaging the worms. In any case, the transferred *Wolbachia* DNA sequences found in the *Wolbachia*-dependent filarial species are most likely degenerate [7] and therefore unable to produce functional proteins.

In this study, we examined two filarial nematode species from different clades [10] that are naturally *Wolbachia*-free, namely *Acanthocheilonema viteae* (a rodent

parasite whose life cycle can be maintained in the laboratory) and *Onchocerca flexuosa* (a parasite of European red deer and a close relative of the agent of African river blindness) [11,12,13]. It has been suggested that the ancestors of these species were colonized in the distant past, as some 90% of filarial nematode species examined to date contain the bacteria [10,11]. We hypothesized that if this is true, HGT may have brought *Wolbachia* DNA into the nuclear genomes of these species prior to endosymbiont loss. We used massively parallel sequencing to survey the genomes of *A. viteae* and *O. flexuosa* in search of *Wolbachia*-like DNA sequences. The presence of *Wolbachia*-like sequences in their nuclear genomes provides the first direct evidence that the ancestors of these species harbored *Wolbachia* endosymbionts that were subsequently lost. Transferred *Wolbachia* genes and/or regulatory elements may help explain the ability of uninfected species to survive without a bacterial partner. Further analysis of transferred genes will provide insight into the nature of the symbiotic relationship between *Wolbachia* and its filarial nematode hosts.

RESULTS

1. Genome sequencing and identification of *Wolbachia*-like sequences

To survey the genomes of *A. viteae* and *O. flexuosa* for transferred *Wolbachia* sequences, fragment and paired-end genomic libraries were sequenced using 454 GS-FLX technology. Two orthologous approaches were undertaken to remove redundancy or capture longer contigs containing *Wolbachia* homologs (see Materials and Methods). *B. malayi*, the only filarial nematode for which draft genome information is available, has an ~95 Mb genome containing ~14% repetitive sequences [14]. Assuming a similar size and structure, we estimate that assembled contigs provide a ~38% coverage of the non-repetitive portions of the *A. viteae* genome. It is not possible to estimate coverage of the *O. flexuosa* genome, as the coverage was insufficient for assembly of paired-end reads. BLASTN analyses identified 45 and 92 genomic fragments containing *Wolbachia*-like sequences in *A. viteae* and *O. flexuosa*, respectively (**Tables 1**, **S1** and **S2**). Subsequent similarity searches found that 14 of the 45 genomic fragments in *A. viteae* and 32 of the 92 genomic fragments in *O. flexuosa* also contain filarial nematode gene homologs. This demonstrates that the *Wolbachia* homologs residing on these fragments are physically integrated into the filarial genomes.

2. Analysis of *Wolbachia*-like sequences

BLAST analysis was used to annotate the genomic DNA fragments identified in this screen (**Tables S1** and **S2**). A total of 49 and 114 *Wolbachia*-like DNA sequences were identified in *A. viteae* and *O. flexuosa*, respectively. The average identity (% ± standard deviation) of the *Wolbachia* homologs to their top BLAST hit was 78±6% in *A. viteae* and 81±6% in *O. flexuosa*, and the average alignment length (bp ± standard deviation)

was 159 ± 83bp and 174 ± 192 bp for *A. viteae* and *O. flexuosa,* respectively. For comparison, the average identity to a filarial nematode gene was 79 ± 16% in *A. viteae* and 83 ± 6% in *O. flexuosa*. Despite low-level sequence coverage, seven *Wolbachia* genes were represented by sequence fragments in both *A. viteae* and *O. flexuosa*. Some of the transferred fragments present in *A. viteae* or *O. flexuosa* also correspond to *Wolbachia* gene fragments present in the nuclear genome of *B. malayi* [7]. So far, none have been identified in all three species (**Table 2**).

3. Cellular processes represented by the transferred DNA

Each of the *Wolbachia*-like gene fragments identified in this study was assigned to a COG functional category in order to determine which cellular processes and pathways were most heavily represented in our transferred fragment collection (**Table 3**). Forty of the 49 *Wolbachia* sequences from *A. viteae* and 104 of the 114 from *O. flexuosa* could be matched to gene from the *Wolbachia* strain wBm from *B. malayi* with a functional role. No COG functional category was identified as over-represented in the HGT sequences as compared to the genome of wBm (Fisher's Exact test, Bonferroni step-down correction, $p<0.01$).

4. Potential source of transferred fragments

Of the 49 *Wolbachia* homologs found in *A. viteae*, 19 (39%) align best to a filarial *Wolbachia* sequence and 30 align best to an insect *Wolbachia* sequence. Likewise, 47 (41%) of the *Wolbachia* homologs in *O. flexuosa* align best to a filarial *Wolbachia* sequence and 67 align best to an insect *Wolbachia* sequence (**Tables S1** and **S2**). Additionally, there was no apparent clustering pattern when transferred sequences were aligned to the circular genome of the *Wolbachia* endosymbiont of *B. malayi* (**Fig. 1**).

Alignment of transferred fragments to the genomes of the *Wolbachia* endosymbionts of *Drosophila melanogaster*, *Drosophila simulans*, and *Culex pipiens* showed a similar lack of clustering. Therefore, we cannot determine whether a large piece of *Wolbachia* DNA (or even an entire bacterial genome) was inserted and subsequently fragmented and scattered in the filarial genomes over time, or if small fragments were shuttled into the genome separately.

5. Mechanism of transfer

The mechanism responsible for DNA transfer from *Wolbachia* is unknown, but the sequence data provide some interesting clues (**Tables S1, S2, S3** and **S4**). *O. flexuosa* contains several homologs to *Wolbachia* proteins involved in bacterial type IV secretion. This suggests that the DNA donor *Wolbachia* strain had a type IV secretion system which could have been capable of shuttling DNA out of the bacterial compartment. Furthermore, remnants of pao retrotransposon sequences were identified in both species (*A. viteae* contig 187 and *O. flexuosa* contig 13). For example, *O. flexuosa* contig 13 contains a *Wolbachia*-like DNA fragment flanked by a pao retrotransposon sequence and a poly(A) tract (**Fig. 2**). Large poly(A) sequences (>20 bases) are present in eight of our reported contigs. Most lie within a few hundred bases of a *Wolbachia*-like sequence. These poly(A) sequences may be remnants of poly(A) tails and suggest retrotransposition of processed eukaryotic mRNAs. Sequence duplication and inverted repeats present in our data are also consistent with transposon insertion sites (**Fig. 2**). Arthropod *Wolbachia* contain prophage sequences thought to enhance DNA rearrangements [15]; one phage sequence was detected in *A. viteae* (wAv16332).

6. Coding potential of transferred fragments

Most of the sequences identified in this screen represent only small portions of *Wolbachia* genes. Some of these sequences, 23 in *A. viteae* and 61 in *O. flexuosa*, are truncated at the end of a contig, so further sequencing will be needed to determine their actual length. The fragments that fall entirely within a sequenced contig (i.e. >25bp from the end of a contig) had average sizes of 146±84 and 183±219 bp in *A. viteae* and *O. flexuosa*, respectively. For comparison, the predicted protein coding genes of the *Wolbachia* endosymbiont of *B. malayi* range from 42 to 2839 amino acids; 35 proteins are predicted to be encoded by sequences shorter than 200 bp (~66 amino acids) (www.ncbi.nlm.nih.gov/sites/entrez?Db=genome&Cmd=Retrieve&dopt=Protein+Table&list_uids=630). BLASTX was able to identify 28 *Wolbachia* homologs in *A. viteae* and 70 in *O. flexuosa*, fewer homologs than the number identified by BLASTN. Ten of the 28 in *A. viteae* and 27 of the 70 homologs in *O. flexuosa* are free of stop codons and frameshift mutations (see **Tables S3** and **S4**).

7. Expression of the transferred fragments

Sybr Green qRT-PCR was used to assess expression of *Wolbachia*-like sequences regardless of the presence of an open reading frame. Stringent controls were used to rule out DNA contamination. Thirty of 42 *A. viteae* and 56 of 87 *O. flexuosa Wolbachia*-like sequences tested were expressed at the RNA level (**Table S5**). 14 of the expressed *Wolbachia* homologs in *A. viteae* and 34 in *O. flexuosa* appear to be transcribed from pseudogenes relative to what is known from the endobacterial genome.

8. Localization of transcripts

In situ RNA hybridization was used to localize two transcripts predicted to arise from pseudogenes of 2-methylthioadenine synthase (2-MAS), which contains a frameshift mutation, and DNA polymerase I (pol I), which contains several premature stop codons. The *Wolbachia* homologs of both of these sequences are involved in nucleic acid synthesis. The 321bp 2-MAS and 415bp polA probe sequences had 83% and 73% identity to their homologs in wBm, respectively. The 2-MAS probe labeled the lateral chords and intrauterine stretched microfilaria, tissues containing *Wolbachia* endobacteria, as well as the intestinal and uterine epithelium in *B. malayi* (**Fig. 3B**). The staining of the intestine and uterine epithelium could be indicative of an expressed HGT fragment in *B. malayi*, as these tissues do not contain *Wolbachia*. In *A. viteae*, this probe labeled developing embryos in females and late spermatogonia in males (**Fig. 3D**, **F**). No staining was seen in the lateral chords. Most of the examined *O. flexuosa* nodule sections did not contain females with developing embryos. However, one *O. flexuosa* nodule containing a young male showed intensive labeling of the testes similar to that seen in male *A. viteae* (**Fig. 3F, H**). The pol I probe weakly labeled *Wolbachia* in the lateral chords (arrows) of *B. malayi* (**Fig. 3J**). Some background staining was observed in the ovaries and uterus using the sense probe (**Fig. 3 I, K**), but this was very weak compared to the strong signal obtained with the anti-sense probe in female *B. malayi* and *A. viteae* (**Fig. 3J, L**). These results show that *Wolbachia*-like pseudogene transcripts can be detected in *Wolbachia*-free filarial species. Expression appears to be tightly regulated, because not all stages, body regions and tissue types were stained.

DISCUSSION

The nuclear genomes of two distantly related, *Wolbachia*-free filarial nematode species contain *Wolbachia*-like DNA sequences that were obtained from a former endosymbiont via HGT. We detected transcription of several *Wolbachia* homologs present in *A. viteae* and *O. flexuosa* despite the fact that many of these sequences are degenerate. *In situ* hybridization showed that two pseudogene transcripts had tissue-specific expression patterns in three filarial species. Our results provide strong evidence to support the hypothesis that the ancestor of extant *Wolbachia*-free filarial species was infected with *Wolbachia*.

Low coverage genome sequencing was used as a cost-effective approach to identify *Wolbachia*-like sequences. Providing draft genomes and a full inventory of all transferred DNA fragments from *A. viteae* and *O. flexuosa* was beyond the scope of this project. Rather, our aim was to provide evidence that the absence of *Wolbachia* in presently uninfected filarial species is due to secondary loss and that DNA had been passed from *Wolbachia* to these species prior to this loss. At present, draft genomes are only available for three parasitic nematode species; over 20 more are in progress [16]. Although priority is still given to pathogens of socio-economic importance, future studies may provide more complete drafts of these genomes at a later date.

The genetic screens described in this report identified 49 and 114 *Wolbachia* homologs in *A. viteae* and *O. flexuosa*, respectively. It will not be possible to determine if *O. flexuosa* and *A. viteae* have the same number of transferred genes until both genomes are fully sequenced. However, it is likely that we have identified more sequences in *O.*

flexuosa due to technical issues like the use of a paired-end sequencing approach as opposed to the traditional fragment library that was used for *A. viteae*.

Several methods are commonly used to detect HGT events [17]. Approaches such as analysis of GC content or codon bias were not feasible because the GC content of filarial worms is similar to that of *Wolbachia* [4], and because we do not have enough information about the genomes of *A. viteae* and *O. flexuosa* to evaluate their codon usage. Our project relied heavily on homology-based analyses. This was ideal for our application because the suspected source of the transfer could be inferred based on the presence of *Wolbachia* in other filarial species and based on the propensity of *Wolbachia* to transfer DNA to the nuclear genomes of its hosts [7].

It is believed that mutualistic *Wolbachia* provide metabolites that are essential for host reproduction, development and survival. The combined genomes of *B. malayi* and its *Wolbachia* endosymbiont encode complete pathways for the biosynthesis of purines, pyrimidines, riboflavin, flavin adenine dinucleotide and heme, pathways that are missing or incomplete in the filarial genome alone [4]. Some of the sequence fragments we identified are homologous to genes involved in these processes (**Tables S1** and **S2**). For instance, several genes related to heme synthesis and export were present in the HGT fragment list. Nematodes, including *C. elegans*, are not known to synthesize their own heme [18], so the presences of a complete heme biosynthetic pathway in either *A. viteae* or *O. flexuosa* would be unique. Since only portions of these genomes have been scanned, further studies are needed to determine the full insert length for many of the known *Wolbachia* homologs and to reveal more transferred sequences, some of which may encode full-length gene products.

The data provide intriguing clues about the donor strain and how transferred DNA was incorporated into the nuclear genome of filarial nematodes. If the transfer came from an extant *Wolbachia* species, the majority of the fragments would show a top BLAST score to this species. We would also expect that the transferred fragments would fall within certain portions of the donor genome if they were incorporated in one or few events. These characteristics were not seen when we compared the transferred fragments to sequenced *Wolbachia* genomes. The variability in the strain of the top BLAST hit and the average percent identity to that hit suggests that transfer(s) either came from a species that is not represented in public databases and/or that the transferred sequences have mutated over time with respect to their parent gene. Either way, these sequence changes must have taken millions of years to accumulate and stabilize in filarial and *Wolbachia* populations [19]. The initial transfer event(s) probably took place between an ancestral *Wolbachia* strain and an ancestral filarial species. It is likely that neither of these exists in the same form today, but we see the "fossilized" evidence of their interaction in the genomes of their descendents.

Questions remain as to how the DNA escaped the bacterial compartment and entered the nucleus of the host cell. It is possible that this may have involved bacterial Type IV secretion, a system conserved in *Wolbachia*. Type IV secretion systems are capable of shuttling proteins and nucleoprotein complexes across membranes and are known to facilitate gene transfer [20,21]. After translocation, DNA may have been inserted into the host genome by transposable elements. Furthermore, the remnants of retrotransposons and large poly(A) stretches found in our sequences suggest that some of

the sequences may be processed pseudogenes derived from mature mRNAs [19]. We have yet to identify the parent copies of these putative processed pseudogenes.

Many of the *Wolbachia* homologs identified in this study show insertions and/or truncations, frameshift mutations and premature stop codons relative to homologs from sequenced *Wolbachia* genomes. Gene fragmentation and degradation suggests a lack of selective pressure to maintain coding capacity, but full-length transcripts and proteins are not always required for biological function. For example, transfer of DNA from mitochondria to the nuclear genome can generate novel exons that alter protein function [22]. Likewise, truncated *Wolbachia* sequences inserted into filarial genes may act as new protein domains that alter or enhance the function of existing nematode proteins. Pseudogenes containing frameshifts and premature stops are abundant in many genomes [23], and widespread transcription of pseudogenes has been reported [24,25]. mRNA recoding mechanisms, which are employed infrequently in most organisms, could allow the translational machinery to produce a protein despite these coding errors [26]. Even if the pseudogene sequences are not translated, recent studies have shown that expressed pseudogenes sequences can regulate expression of other genes through RNA interference [27,28,29,30]. Further studies will determine which of these mechanisms might allow the transferred *Wolbachia* sequences reported here to contribute to filarial biology.

We have provided strong evidence that two distantly related *Wolbachia*-free filarial nematodes contain *Wolbachia*-like DNA in their nuclear genomes. HGT from bacteria may be a relatively common phenomenon in nematode phylogeny. As previously mentioned, *Wolbachia* have inserted DNA in the nuclear genomes of endosymbiont-dependent filarial nematode hosts [7]. In these cases, the source of the transfer is obvious

because the mutualistic relationship between the two organisms has been maintained, but the functions of the inserted sequences are still unknown. The converse situation exists in plant parasitic nematodes. Cellulase enzymes that support plant parasitism were probably obtained from bacteria via HGT [31,32]. These cellulase genes show a high degree of homology to genes from a wide range of bacterial species. The exact source of the transferred DNA is difficult to pinpoint because the relationship between the DNA donor and recipient has not been maintained. The case of *A. viteae* and *O. flexuosa* is special because the relationship with the DNA donor has not been maintained, but the donor can be easily identified by BLAST homology and by the presence of *Wolbachia* in most filarial species. Future studies will generate a comprehensive list of transferred fragments in several *Wolbachia*-free filarial species. This information may identify key genes and pathways that explain the fascinating and medically important symbiotic relationship between filarial worms and *Wolbachia* endobacteria.

MATERIALS AND METHODS

1. Parasite material and DNA isolation

Adult *O. flexuosa* worms were collected from red deer (*Cervus elaphus*) in Germany (Schleswig-Holstein). Adult worms were dissected from nodules collected from freshly shot deer [13]. Adult *A. viteae* and *B. malayi* were obtained from experimentally infected Mongolian gerbils as previously described [33,34]. DNA was isolated from adult worms using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA).

2. Library construction, sequencing and BLAST analysis

A fragment library was created from 15 μg of *A. viteae* DNA. Two runs on a Genome Sequencer FLX (454 Life Sciences/Roche Diagnostics, Branford, CT) using standard FLX chemistry generated 768,909 reads and a total of 181 Mbp of sequence. Sff files were deposited in the NCBI short read archive (SRX001994). Newbler 1.1.03.24 assembled reads into 68,805 contigs containing ~31 Mbp of sequence using default parameters. The largest contig was 6.7 kbp and the average size was 0.9 kbp. Contigs were analyzed by BLASTN against the Genbank nucleotide collection (nt) and by BLASTX against the Genbank non-redundant protein database (nr) using NCBI BLASTALL to identify regions with a top hit to *Wolbachia* with an e-value less than 1 x 10^{-5}. This cutoff was used for all BLAST searches. All *A. viteae* contigs larger than 600 bp were also split into smaller fragments (optimal maximum and minimum were set to 600 and 300 bp, respectively) using a custom Perl script and re-analyzed by BLASTN (nt_05292009) and BLASTX (nr_05292009) to ensure that all *Wolbachia* homologs were identified despite strong homology to different genes/proteins in other regions of the sequence. Contigs of interest were PCR amplified and cloned using the TOPO TA

Cloning Kit for Sequencing or the TOPO-XL PCR Cloning Kit (Invitrogen, Carlsbad, CA, USA), and sequenced for verification. Sequences are posted on the Whole Genome Shotgun FTP site on Nematode.net (www.nematode.net/FTP/index.php).

O. flexuosa paired-end genomic DNA libraries were constructed as previously described [35] using 5.0 μg of DNA. One run on a Genome Sequencer FLX using standard FLX chemistry produced 516,745 reads containing 135 Mbp of sequence including adapter sequences. Sff files were deposited in the NCBI short read archive (SRX015550). A custom Perl script was designed for removal of the 44 bp adapter and separation of the paired end (PE) sequences which were then analyzed by BLASTX (nr_05292009) and BLASTN (nt_05292009). Complete, unsorted reads were also analyzed by BLASTX (nr_09182007) and BLASTN (nt_04062009) using WU-BLAST 2.0. 137 reads containing *Wolbachia*-like sequences were identified. Read sequences were used to amplify larger portions of the *O. flexuosa* genome. After sequencing the larger genomic fragments, redundant/overlapping fragments were collapsed using Contig Express (Invitrogen). This produced 42 large fragments (those amplified using primers matching to PE sequences) ranging in size from 1124 to 7725 bp with an average size of 3028±1386 bp and 50 smaller sequences (those representing assembled or unassembled portions of a PE read) ranging in size from 55 to 548 bp with an average size of 243±83 bp. All fragments sequences are posted on the Data FTP page of Nematode.net.

3. Assessment of COG functional roles and mapping to the genome of the *Wolbachia* endosymbiont of *B. malayi*

Wolbachia-like sequences were compared against the genome of the *Wolbachia* endosymbiont strain TRS of *B. malayi* (wBm) using BLASTN, and the locus tag of the

top hit was recorded. The *Wolbachia*-like sequence was assigned to the same COG functional category as its wBm homolog as reported in NCBI Entrez. The Open Source Python (http://www.python.org/) library ReportLab (http://www.reportlab.com/software/opensource/rl-toolkit/) was used to generate a figure that marks the position of the homologous locus. Tick marks represent the midpoint of the coordinates for each locus tag as extracted from the GenBank data file for the wBm genome (NC_006833).

4. RNA isolation, cDNA synthesis and qRT-PCR

Adult worms were homogenized in 1mL Trizol (Invitrogen), and RNA was isolated using organic extraction followed by column purification using an RNeasy Mini Kit (Qiagen) including the on-column DNase digest. A second DNase treatment was performed using Ambion's DNA-*free* DNase kit (Applied Biosystems, Austin, TX, USA). rRNA was depleted using the RiboMinus Eukaryote Kit for RNA-Seq (Invitrogen). Samples were tested for DNA contamination by conventional PCR using 35 cycles to ensure that no products were amplified. cDNA was made from 500ng mRNA using qScript cDNA super mix (Quanta Biosciences, Gaithersburg, MD, USA). Conventional PCR assays using primers designed to span introns were used to test for DNA contamination again. Intron-spanning primer sequences are provided in **Table S6**. These target sequences from *O. flexuosa* with homology to hypothetical proteins in *B. malayi* that were readily amplifiable from *O. flexuosa* genomic and cDNA. Only Of_IC 1 worked well with *A. viteae* mRNA, which was also rigorously tested with wAv primers used in this study prior to cDNA synthesis (see **Table S6**). Specific *Wolbachia* sequences were detected in cDNA by SybrGreen qRT-PCR using PerfeCTa SYBR Green FastMix,

ROX (Quanta Biosciences). Primer sequences are reported in **Table S6**. Reactions were done in duplicate to ensure accuracy, and all qRT-PCR experiments included a 10 ng DNA positive control, 10 – 0.1 ng ten-fold dilutions of cDNA, and a 10 ng RNA negative control. Dissociation curves were examined to rule out non-specific amplification, and all products were assessed by agarose gel electrophoresis. Sequences were reported as expressed at the transcript level when signal for the 0.1ng cDNA template crossed the cycle threshold at least 3 cycles before the mRNA negative control.

5. *In situ* hybridization

Adult *B. malayi* and *A. viteae* worms and *O. flexuosa* nodules were fixed for 24-72 h in DEPC-treated 4% buffered formaldehyde and embedded in paraffin using standard histological procedures. Two sequences were amplified from *O. flexuosa* cDNA and cloned into a dual promoter PCRII plasmid (Invitrogen). Primer sequences are reported in **Table S6**. After linearization of the plasmid, biotinylated antisense probes and sense negative controls were prepared with Ambion Megascript T7 and Sp6 high yield transcription kits (Applied Biosystems). Following DNase digestion (Roche, Indianapolis, IN, USA), probes were concentrated by ethanol precipitation, re-suspended in DEPC-treated water, and stored at -20°C. Sections (5 μm) were deparaffinized, digested with pepsin HCl for 7 min, and hybridized at 37°C overnight in a humid chamber with 1μg of RNA probe in hybridization buffer (50% formamide, 5XSSC, 0.3 mg/ml yeast tRNA, 100 μg/ml heparin, 1X Denhardts Solution, 0.1% Tween 20, 0.1% CHAPS and 5mM EDTA). A stringency wash was performed at 60°C for 30 min, and detection was performed using the 'In situ Hybridization Detection System' (K0601, DakoCytomation, Hamburg, Germany). Sections were incubated for 20 min with

streptavidin-AP conjugate at room temperature. BCIP/NBT substrate solution was used for 10-30 min to visualize the RNA target. Sections were examined using an Olympus-BX40 microscope (Olympus, Tokyo, Japan) and photographed with an Olympus DP70 microscope digital camera.

ACKNOWLEDGEMENTS

We would like to thank Ramakrishna U. Rao, Washington University and Richard Lucius, Humboldt University Berlin, for providing *B. malayi* and *A. viteae* material. Anita Plenge-Boenig, Insa Bonow, Frank Geisinger and the late Manfred Kroemer, all from the Bernhard Nocht Institute, were instrumental in the collection *O. flexuosa* material. We also thank Andrew Mullin for his expert technical assistance and Nathan McNulty for bioinformatic advice.

REFERENCES

1. Taylor MJ, Bandi C, Hoerauf A (2005) *Wolbachia* bacterial endosymbionts of filarial nematodes. Adv Parasitol 60: 245-284.
2. Hoerauf A, Mand S, Fischer K, Kruppa T, Marfo-Debrekyei Y, et al. (2003) Doxycycline as a novel strategy against bancroftian filariasis-depletion of *Wolbachia* endosymbionts from *Wuchereria bancrofti* and stop of microfilaria production. Med Microbiol Immunol 192: 211-216.
3. Hoerauf A, Nissen-Pahle K, Schmetz C, Henkle-Duhrsen K, Blaxter ML, et al. (1999) Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility. J Clin Invest 103: 11-18.
4. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, et al. (2005) The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. PLoS Biol 3: e121.
5. Kozek WJ (1977) Transovarially-transmitted intracellular microorganisms in adult and larval stages of *Brugia malayi*. J Parasitol 63: 992-1000.
6. Fenn K, Conlon C, Jones M, Quail MA, Holroyd NE, et al. (2006) Phylogenetic relationships of the *Wolbachia* of nematodes and arthropods. PLoS Pathog 2: e94.
7. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753-1756.
8. Kondo N, Nikoh N, Ijichi N, Shimada M, Fukatsu T (2002) Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci U S A 99: 14280-14285.
9. Nikoh N, Tanaka K, Shibata F, Kondo N, Hizume M, et al. (2007) *Wolbachia* genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. Genome Res.
10. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. Int J Parasitol 34: 191-203.
11. Bandi C, Anderson TJ, Genchi C, Blaxter ML (1998) Phylogeny of *Wolbachia* in filarial nematodes. Proc Biol Sci 265: 2407-2413.
12. McGarry HF, Pfarr K, Egerton G, Hoerauf A, Akue JP, et al. (2003) Evidence against *Wolbachia* symbiosis in *Loa loa*. Filaria J 2: 9.
13. Plenge-Bonig A, Kromer M, Buttner DW (1995) Light and electron microscopy studies on *Onchocerca jakutensis* and *O. flexuosa* of red deer show different host-parasite interactions. Parasitol Res 81: 66-73.
14. Scott AL, Ghedin E (2009) The genome of *Brugia malayi* - all worms are not created equal. Parasitol Int 58: 6-11.
15. Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, et al. (2004) Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. PLoS Biol 2: E69.
16. Brindley PJ, Mitreva M, Ghedin E, Lustigman S (2009) Helminth genomics: the implications for human health. PLoS Negl Trop Dis 3: e538.

17. Whitaker JW, McConkey GA, Westhead DR (2009) Prediction of horizontal gene transfers in eukaryotes: approaches and challenges. Biochem Soc Trans 37: 792-795.
18. Rao AU, Carta LK, Lesuisse E, Hamza I (2005) Lack of heme synthesis in a free-living eukaryote. Proc Natl Acad Sci U S A 102: 4270-4275.
19. Podlaha O, Zhang J (2009) Processed pseudogenes: the 'fossilized footprints' of past gene expression. Trends Genet.
20. Juhas M, Crook DW, Hood DW (2008) Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. Cell Microbiol 10: 2377-2386.
21. Pichon S, Bouchon D, Cordaux R, Chen L, Garrett RA, et al. (2009) Conservation of the Type IV secretion system throughout *Wolbachia* evolution. Biochem Biophys Res Commun 385: 557-562.
22. Noutsos C, Kleine T, Armbruster U, DalCorso G, Leister D (2007) Nuclear insertions of organellar DNA can create novel patches of functional exon sequences. Trends Genet 23: 597-601.
23. Harrison PM, Gerstein M (2002) Studying genomes through the aeons: protein families, pseudogenes and proteome evolution. J Mol Biol 318: 1155-1174.
24. Williams DL, Slayden RA, Amin A, Martinez AN, Pittman TL, et al. (2009) Implications of high level pseudogene transcription in *Mycobacterium leprae*. BMC Genomics 10: 397.
25. Zheng D, Frankish A, Baertsch R, Kapranov P, Reymond A, et al. (2007) Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. Genome Res 17: 839-851.
26. Baranov PV, Gurvich OL, Fayet O, Prere MF, Miller WA, et al. (2001) RECODE: a database of frameshifting, bypassing and codon redefinition utilized for gene expression. Nucleic Acids Res 29: 264-267.
27. Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, et al. (2008) The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453: 803-806.
28. Sasidharan R, Gerstein M (2008) Genomics: protein fossils live on as RNA. Nature 453: 729-731.
29. Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, et al. (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453: 534-538.
30. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, et al. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453: 539-543.
31. Bird DM, Opperman CH, Davies KG (2003) Interactions between bacteria and plant-parasitic nematodes: now and then. Int J Parasitol 33: 1269-1276.
32. Mitreva M, Smant G, Helder J (2009) Role of horizontal gene transfer in the evolution of plant parasitism among nematodes. Methods Mol Biol 532: 517-535.
33. Ash LR, Riley JM (1970) Development of subperiodic *Brugia malayi* in the jird, *Meriones unguiculatus*, with notes on infections in other rodents. J Parasitol 56: 969-973.

34. Lucius R, Textor G (1995) *Acanthocheilonema viteae*: rational design of the life cycle to increase production of parasite material using less experimental animals. Appl Parasitol 36: 22-33.
35. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, et al. (2007) Paired-end mapping reveals extensive structural variation in the human genome. Science 318: 420-426.

FIGURES

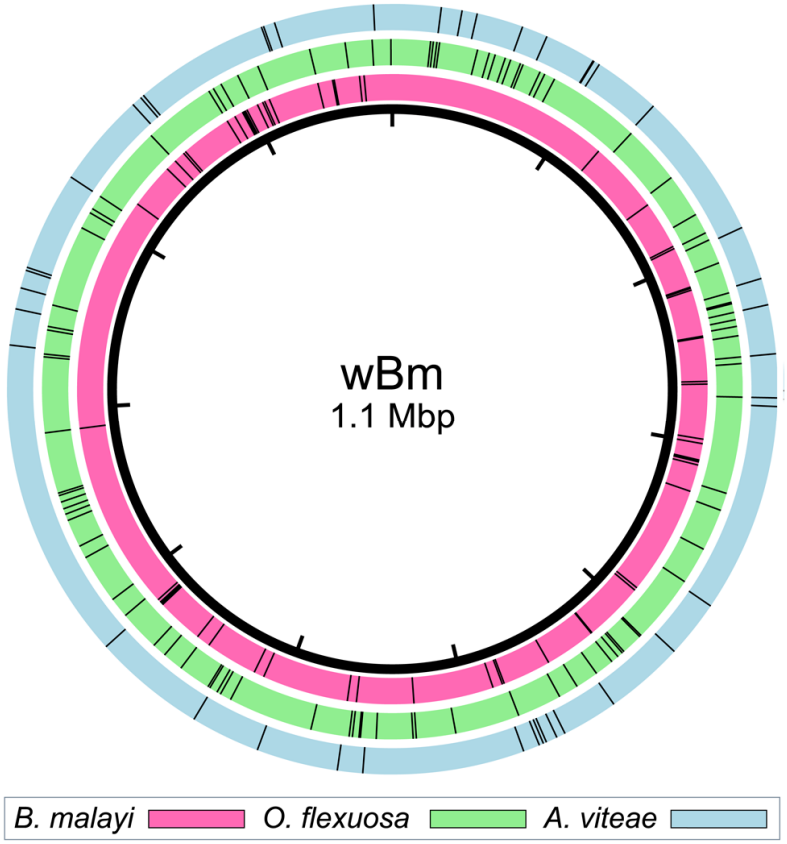


Figure 1: Mapping of transferred fragments to a sequenced *Wolbachia* genome. Black circle represents the 1.1 Mbp genome of the *Wolbachia* endosymbiont of *B. malayi*. Tick marks in the colored outer rings indicate where a transferred DNA fragment found in the indicated species would align to the *Wolbachia* genome. Fragments found in the *B. malayi* genome were previously described by Dunning Hotopp et al. [7].

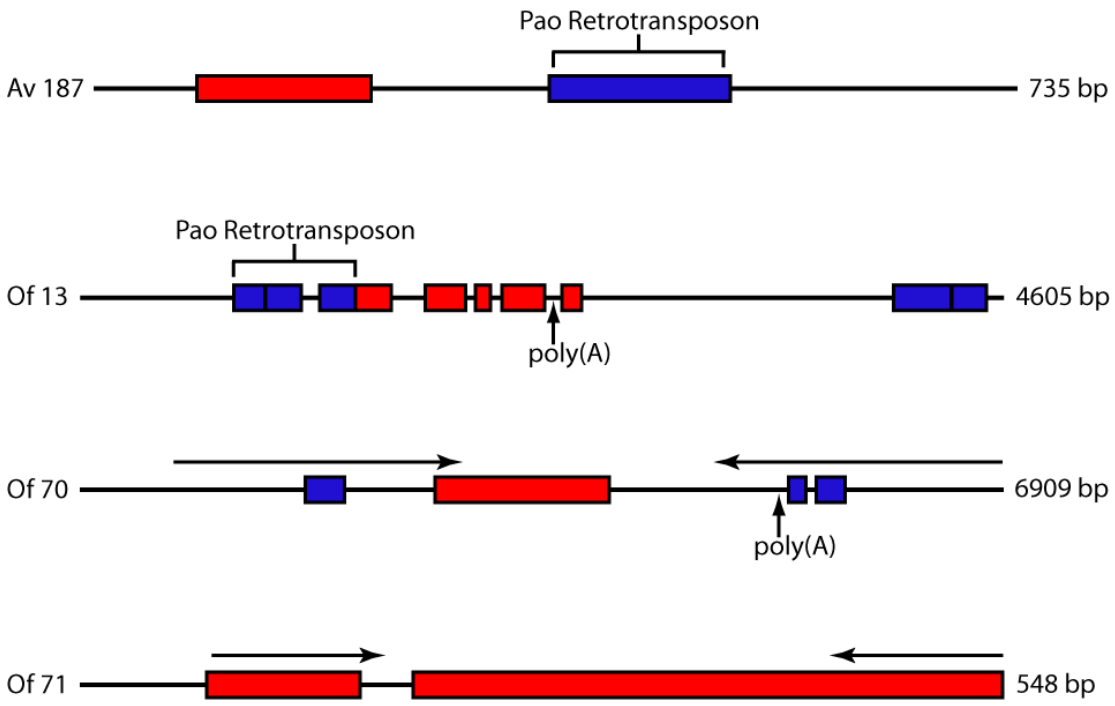


Figure 2: Schematic of genomic DNA fragments containing *Wolbachia* homologs. The exact annotation and coordinates of each of the homologs depicted can be found in **Tables S1** and **S2**. Blue blocks represent regions of homology to nematode sequences while red blocks represent regions homologous to *Wolbachia* sequences. Horizontal arrows represent inverted repeats in the DNA sequence. Inverted repeat segments in Of70 share 88% identity with one another while the repeated segments in Of71 share 82.8% identity.

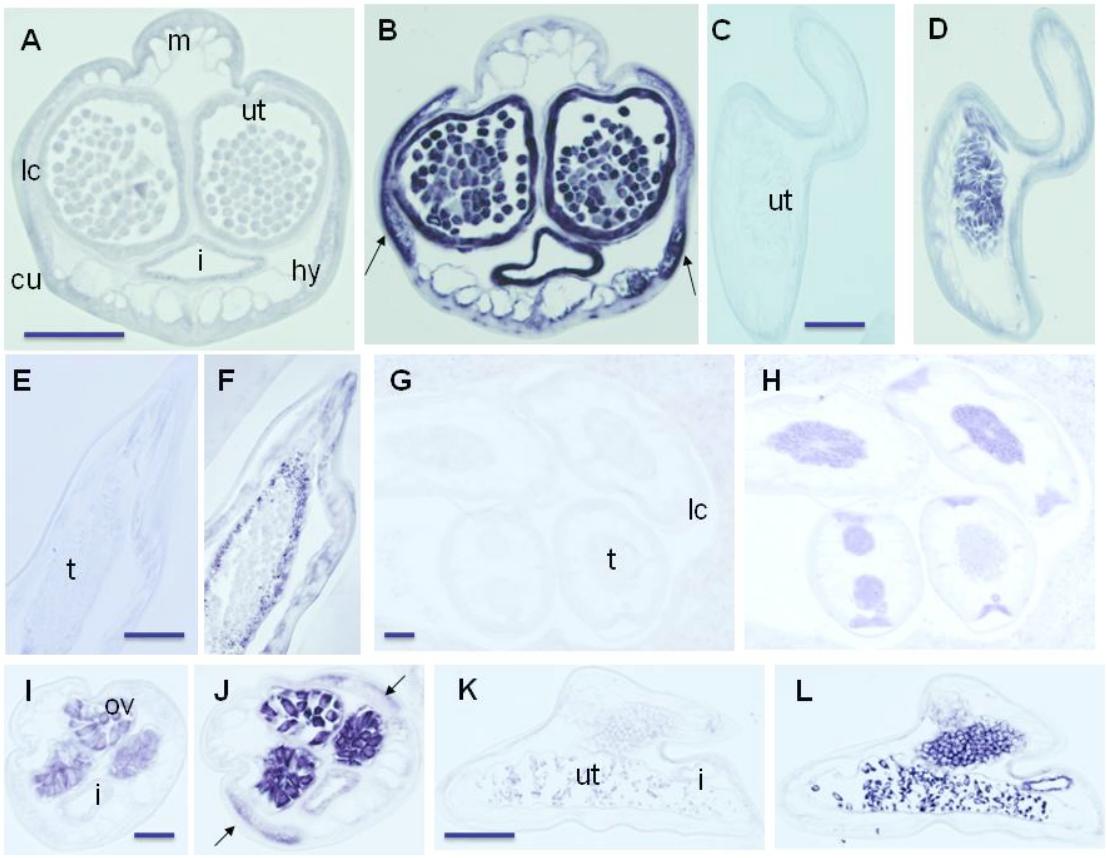


Figure 3: *In situ* hybridization of adult filarial worms. A-H are stained with probes made from an *O. flexuosa* sequence with homology to *Wolbachia* 2- methylthioadenine synthase (2-MAS, wOf53). A, C, E and G are stained with the sense probe (negative control), while B, D, F and H show matching consecutive sections stained with the antisense probes. 2-MAS probe labels lateral cords, intrauterine stretched microfilaria and uterine and intestinal epithelium of female *B. malayi* (B), the oocytes and uterus of female *A. viteae* (D), the spermatogonia in the periphery of the testes of male *A. viteae* (F), and lateral chords and different stages of sperm development in the testes of male *O. flexuosa* (H). I-L are stained with probes made from an *O. flexuosa* sequence with homology to *Wolbachia* DNA polymerase I (pol I, wOf88). I and K are stained with sense probes (negative control) while H and L are the matching consecutive sections stained with antisense probes. pol I probe labels ovaries and granular structures resembling *Wolbachia* (arrows) in the lateral chords of female *B. malayi* (J) and oocytes in female *A. viteae* (L). Abbreviations: m, musculature; i, intestine; lc, lateral chord; cu, cuticle; hy, hypodermis; t, testes; ut, uterus; ov, ovary. Scale bar 40 μm.

TABLES

Table 1: Identification of *Wolbachia*-like sequences.

Species	**Library Setup**	**Fragments with *Wolbachia*-like sequences**	***Wolbachia* Homologs**	**Average %ID of *Wolbachia* homologs**	**Average length of *Wolbachia* homologs**	**Fragments with Junctions***
O. flexuosa	Paired end	92	114	78±6%	158.9±82.6 bp	32
A. viteae	Fragment	45	49	81±6%	173.6±191.8 bp	14

*Fragments with junctions are defined as continuous pieces of DNA that contain sequences homologous to both *Wolbachia* and nematode genes.

Table 2: *Wolbachia* homologs found in multiple species. *Wolbachia* inserts in the nuclear genome of *B. malayi* were described previously [7]. Shared homologs were identified by alignment to *Wolbachia* sequences with the same locus tag by BLASTN. Presence of sequences in the nuclear genome or among transcripts (+), lack of expression at RNA level (-), and an inability to test for expression (n/a) are noted.

	A. viteae		***O. flexuosa***		***B. malayi***
Annotation	**DNA**	**RNA**	**DNA**	**RNA**	**DNA**
rod shape-determining protein RodA	+	+	+	-	
4-Hydroxy-3-methylbut-2-enyl diphosphate reductase, IspH	+	-	+	+	
ATP-binding subunit of Clp protease and DnaK/DnaJ chaperones	+	+	+	+	
methionyl-tRNA synthetase	+	-	+	+	
ribosomal large subunit pseudouridine synthase C, putative	+	+	+	+	
ribosomal protein L27	+	+	+	n/a	
DNA-directed RNA polymerase, beta/beta' subunits	+	+	+	-	
valyl-tRNA synthetase			+	+	+
type IV secretion system protein VirB4, putative			+	+	+
ATP-dependent Zn protease, HflB			+	-	+
dimethyladenosine transferase			+	-	+
DNA polymerase III, beta subunit			+	n/a	+
cell cycle protein (ftsZ) gene			+	+	+
DNA polymerase III, gamma/tau subunit	+	+			+
IMP dehydrogenase, GuaB	+	+			+

Table 3: Assignment of *Wolbachia*-like sequences to COG functional categories. *Wolbachia* homologs in *A. viteae* and *O. flexuosa* were identified based on BLAST homology. The *Wolbachia*-like sequences were assigned to the same category their wBm orthologs.

Description	*O. flexuosa*		*A. viteae*		*wBm*	
	# loci	% of total	# loci	% of total	# loci	% of total
Information Storage and Processing						
Translation	15	13.2%	10	20.4%	121	15.0%
Transcription	3	2.6%	2	4.1%	18	2.2%
Replication, recombination and repair	7	6.1%	6	12.2%	54	6.7%
Cellular Processes						
Cell cycle control, mitosis and meiosis	3	2.6%	2	4.1%	9	1.1%
Defense mechanisms	0	0.0%	0	0.0%	2	0.2%
Signal transduction mechanisms	2	1.8%	0	0.0%	10	1.2%
Cell wall/membrane biogenesis	6	5.3%	3	6.1%	33	4.1%
Cell motility	0	0.0%	0	0.0%	1	0.1%
Intracellular trafficking and secretion	7	6.1%	0	0.0%	29	3.6%
Posttranslational modification, protein turnover, chaperones	8	7.0%	1	2.0%	51	6.3%
Metabolism						
Energy production and conversion	10	8.8%	3	6.1%	68	8.4%
Carbohydrate transport and metabolism	1	0.9%	3	6.1%	24	3.0%
Amino acid transport and metabolism	10	8.8%	5	10.2%	38	4.7%
Nucleotide transport and metabolism	8	7.0%	2	4.1%	37	4.6%
Coenzyme transport and metabolism	2	1.8%	3	6.1%	33	4.1%
Lipid transport and metabolism	5	4.4%	2	4.1%	26	3.2%
Inorganic ion transport and metabolism	5	4.4%	0	0.0%	35	4.3%
Secondary metabolites biosynthesis, transport and catabolism	1	0.9%	0	0.0%	11	1.4%
Poorly Characterized						
General function prediction only	6	5.3%	3	6.1%	63	7.8%
Function unknown	5	4.4%	1	2.0%	31	3.9%
Not in COGs	8	7.0%	1	2.0%	176	21.9%

Total	**114**		**49**		**805**	

See the supplementary data CD for the following tables:

Table S1: BLASTN annotation of *Acanthocheilonema viteae* genomic DNA fragments.

Table S2: BLASTN annotation of *Onchocerca flexuosa* genomic DNA fragments.

Table S3: BLASTX annotation of *Acanthocheilonema viteae* genomic DNA fragments.

Table S4: BLASTX annotation of *Onchocerca flexuosa* genomic DNA fragments.

Table S5: Results of qRT-PCR expression studies and presence of potential open reading frames.

Table S6: SYBR green qRT-PCR and RNA in situ hybridization probe primers.

CHAPTER 4:

The Transcriptome and Proteome of a *Wolbachia*-Free Filarial Parasite, *Onchocerca flexuosa*, Provide Evidence of Trans-Domain Horizontal Gene Transfer

In preparation for submission to PLoS Biology.

Samantha N. McNulty[1], Sahar Abubucker[2], Gabriel M. Simon[3], Makedonka Mitreva[2], Nathan P. McNulty[3], Kerstin Fischer[1], Kurt C. Curtis[1], Norbert W. Brattig[4], Gary J. Weil[1], Peter U. Fischer[1]

[1] Infectious Diseases Division, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA

[2]The Genome Institute, Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA

[3]Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA

[4] Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany

PREFACE

SNM prepared *Onchocerca flexuosa* RNA, performed bioinformatic analyses of assembled transcriptomic and proteomic data, designed *in situ* probes, and wrote the first complete draft of this manuscript. Comments from co-authors were incorporated into the final version presented here.

ABSTRACT

While many filarial nematode species depend on *Wolbachia* endobacteria to carry out their lifecycle, others are naturally *Wolbachia*-free. Comparing the gene content and expression of closely related *Wolbachia*-dependent and independent filarial nematodes may help to explain this paradox. To address the relative paucity of molecular information on *Wolbachia*-free filarial parasites, we performed a tandem analysis of the transcriptome and proteome of the *Wolbachia*-free species *Onchocerca flexuosa*. Roche/454 sequencing of an *O. flexuosa* cDNA library produced 16,814 isogroups and 47,252 singletons. 1,803 proteins were identified from worm lysate by mass spectrometry using a comparative database containing transcript translations and proteins from related species. Computationally predicted and experimentally verified proteins were classified at the domain and pathway levels. Many *Wolbachia*-like sequences were found among the transcripts, and two *Wolbachia*-like proteins were detected by mass spectrometry. Antibodies against one of the *Wolbachia*-like proteins labeled a single 38kDa band on Western blots of *O. flexuosa* lysate and labeled specific worm tissues by immunohistology. Our study has expanded the volume of sequence and proteomic data available from *O. flexuosa* and has provided the first evidence of protein-level expression of a *Wolbachia*-like gene in a *Wolbachia*-free filarial nematode. Future studies will be required to determine the exact functions of *Wolbachia*-like proteins in *O. flexuosa* and to assess their roles in worm biology.

INTRODUCTION

Filarial nematodes are medically and economically important parasites that infect humans and animals across the globe. Due to their significant impact on human health, a global effort is underway to eliminate lymphatic filariaisis (LF) and to control onchocerciasis [1,2]. The anthelminthic drugs employed in these programs kill larval parasites and temporarily sterilize adult worms. Therefore, periodic treatments must be repeated over the reproductive lifespan of the worms (>5 years for lymphatic filariasis and >10 years for onchocerchiasis) in order to eliminate the infection from endemic areas. Macrofilaricidal drugs would greatly increase the efficiency of parasite eradication.

Wolbachia endobacteria are necessary for the survival and reproduction of many filarial species [3,4] and are a practical target for antifilarial drugs [5]. Antibiotics that are effective against *Wolbachia* (e.g., tetracycline or doxycycline) can be used to treat human filarial infections like LF and onchocerciasis [3,6,7,8], but they are not ideal for mass drug administration regimes because of the long treatment courses required and contraindications to their use in children. Antibiotics designed to impact the specific bacterial pathways or processes necessary for worm survival might be more effective. Unfortunately, the mechanistic details of the worm-bacteria interdependence that might pinpoint appropriate target pathways or processes are poorly understood.

Not all filarial nematodes are *Wolbachia*-dependent [9,10]. Clearly, these *Wolbachia*-free species must possess some "enhanced" function(s) in relation to their *Wolbachia*-dependent counterparts. The genomes and proteomes of *Brugia malayi* and its *Wolbachia* endosymbiont have provided some insight into the processes mediated by each partner in this symbiotic relationship [11,12,13]. It would be informative to compare this

information with similar data collected from closely related, *Wolbachia*-free counterpart. Genes encoded by *Wolbachia* and by *Wolbachia*-free worms but not by *Wolbachia*-dependent worms may be critical to understanding the mutualistic interaction(s) between these obligate endosymbionts and their worm hosts.

Horizontal gene transfer (HGT) from *Wolbachia* to the invertebrate host is a well-known driver of genome evolution in insects and *Wolbachia*-dependent filarial nematodes [14]. Genomic surveys of two *Wolbachia*-free species *Acanthocheilonema viteae* and *Onchocerca flexuosa*, uncovered evidence of HGT from *Wolbachia*, indicating that these species must have been colonized in the distant past [15]. Transferred *Wolbachia* genes or gene fragments may contribute critical functions that allow these worms to live without endosymbionts. However, the extent and consequences of this trans-domain HGT have not been fully explored.

Here we report the results of a survey of the transcriptome and proteome of *O. flexuosa*, a *Wolbachia*-free parasite of European red deer [16]. Among the 16,814 isogroups and 47,252 singleton transcripts identified in this study, 62 isogroups and 51 singletons showed homology to *Wolbachia* proteins. A total of 1,803 proteins were identified in our shotgun proteomic analysis; two of these may be derived from horizontally transferred *Wolbachia* genes. Polyclonal antibodies raised against one of the *Wolbachia*-like peptides label a single 38 kDa band on Western blots of adult worm lysate and stain specific worm tissues by immunohistology. Future studies will be required to determine the exact role of this and other *Wolbachia*-like proteins in *Wolbachia*-free filarial nematodes.

RESULTS

Transcriptome sequencing, assembly, and quality assessment

A total of 1,334,767 reads were generated from an adult stage, poly(A) selected, normalized *O. flexuosa* cDNA library by Roche/454 sequencing and assembled with 2,124 conventional ESTs that were already available in the NCBI EST database. A cDNA-specific protocol was used to assemble sequences in order to account for alternative splicing. The hybrid assembly yielded 16,814 isogroups (putative genes) containing 25,222 isotigs (alternatively spliced isoforms) and 47,252 singletons (**Table 1**). Isotig sequences were deposited in the GenBank transcriptome shotgun assembly database under BioProject number 62565 (accession numbers JI459010-JI484230). All sequences used in subsequent analyses, including singletons, are available at Nematode.net [17].

Using the single-copy conserved eukaryotic genes (CEGs) as a reference [18,19], the completeness of the *O. flexuosa* transcriptome was estimated at approximately 41%, as 101 of the 248 CEG's were identified from the isotigs and singletons. Fragmentation, defined as non-overlapping contigs representing the same gene/transcript, was estimated at 32% or 25% by comparison with *B. malayi* and *C. elegans* protein sequences, respectively.

23.5% of the isogroups derived from the *O. flexuosa* transcriptome contained more than one isotig, an indication that alternative splicing (AS) is common in this parasite. Among AS isogroups, there was an average of 3.1 isotigs per isogroup. For comparison, nearly 25% of *C. elegans* genes are thought to undergo some form of AS (e.g., tissue, sex or stage-specific) with an average of two isoforms per AS gene [20].

Annotation of the *O. flexuosa* transcriptome

59.6% of the isogroups (67.2% of isotigs) and 34.6% of the singletons shared homology with known proteins (**Table 2**). 6,619 of the 16,814 isogroups (39.4%) were matched to homologs in the non-redundant protein database (nr) by BLASTX. Not surprisingly, more than 95% of these matches were to proteins from *B. malayi* and *C. elegans*. 32.0% of the 21,252 *B. malayi* protein sequences and 11.0% of the 24,599 *C. elegans* protein sequences available from WormBase have putative orthologs in our *O. flexuosa* transcriptome (**Table 2**). The average percent identity between high scoring segment pairs was 80.7% with *B. malayi* proteins and 70.0% with *C. elegans* proteins.

A total of 68,402 unique peptide translations (available at Nematode.net) were obtained from 16,807 isogroups (25,205 isotigs) and 47,252 singletons (**Table 1**). Translations from 3,853 isogroups and 2,121 singletons share sequence similarity with 2,804 unique InterPro domains. The most abundant domains were collagen triple helix repeats, kinase domains, and WD40 related domains (**Table 3**).

Translations from 5,159 isogroups and singletons were assigned to 2,049 unique KEGG orthologous (KO) groups, which were further binned into pathway categories (**Table 4**) and modules (**Table S1**). “Folding, sorting and degradation”, “transcription, replication and repair”, and “signal transduction” were among the most heavily represented categories in this analysis. Genes related to the de novo synthesis of riboflavin, heme and nucleotides are missing from the nuclear genome of *B. malayi* but present in the genome of its *Wolbachia* endosymbiont. Therefore, it has been suggested that *B. malayi* may rely on *Wolbachia* as a source of these substances. We identified pathway modules related to these processes (highlighted in **Table 5)**. More sequences

related to uridine monophosphate and inosine monophosphate biosynthesis (required for the de novo synthesis of purines and pyrimidines, respectively) were identified from *O. flexuosa* transcripts than from the nearly complete genome of *B. malayi*, although several enzymes involved in these pathways have yet to be identified.

Analysis of the *O. flexuosa* proteome

Adult worm lysate was subjected to shotgun proteomic analysis in order to characterize the *O. flexuosa* proteome. Using this method, peptides are identified by matching experimentally obtained mass spectra to protein sequences from the organism of interest. The *O. flexuosa* genome has not been sequenced, so we compiled a custom database for comparison with mass spectrometry (MS) data. Our database contained the peptide translations from the *O. flexuosa* transcriptome and all known proteins related to the phylum Nematoda and the genus *Wolbachia*. The interrogation of this database resulted in the successful characterization of 8,537 unique peptides that mapped to 2,685 database sequences. Matches were further grouped into 1,803 proteins (**Table S2**). Of the 1,803 proteins, 673 matched *O. flexuosa* peptide translations, 1,077 matched homologous sequences from other organisms, and 53 matched both. As expected, sequences from *Loa loa* and *B. malayi* were the most informative in identifying *O. flexuosa* peptides (**Fig. 1**). Database entries identified by MS are available at Nematode.net [17].

1,573 of the 1,803 protein groups identified by MS share homology with 1,516 InterPro domains. Thioredoxin folds, concanavalin A-like lectin/glucanase domains and immunoglobulin fold domains were among the most abundant (**Table 3**). 608 proteins were associated with 446 unique KO groups, which were binned into pathway categories

(**Table 4**) and modules (**Table S1**). "Folding, sorting and degradation", "translation", "energy metabolism" and "carbohydrate metabolism" were among the most heavily represented categories. Little representation was seen in the highlighted pathway modules (**Table 5**).

Horizontal gene transfer from *Wolbachia* into the *O. flexuosa* genome

Regions of 62 isogroups (73 isotigs) and 36 singletons share homology with *Wolbachia* proteins (**Table 2** and **Table S3**). 17 of these isogroups (23 isotigs) and 3 of the singletons also contain regions homologous to nematode genes. The *Wolbachia*-like sequences range in size from 62-308 bp and represent fragments of 97 unique *Wolbachia* proteins. 43 of the *Wolbachia* genes with homologs in the transcriptome are from filarial *Wolbachia* strains while the remaining 54 are from insect *Wolbachia* strains. 11 of the 97 *Wolbachia* genes with homologs in the *O. flexuosa* transcriptome have homologs in previously reported *O. flexuosa* genomic sequences; qRT-PCR studies had already shown that 7 of these sequences are expressed at the RNA level [15].

While many of the *Wolbachia*-like sequences contain stop codons and frameshift mutations, 62 peptide translations from 34 isogroups (36 isotigs) and 27 singletons show homology to *Wolbachia* proteins. These translations range in size from 38-216 aa with an average of 65.8±37.3 aa and appear to have diverse functions. The *Wolbachia*-like peptides from 58 isogroups and singletons could be assigned to 41 unique KO groups; the KO groups were binned into 26 pathway modules, all of which are sparsely populated (**Table S4**). Represented pathway modules include inosine monophosphate biosynthesis, uridine monophosphate biosynthesis and heme biosynthesis (**Table 5**).

Although none of the *Wolbachia*-like peptides predicted from the *O. flexuosa* transcriptome were identified by MS, experimentally detected peptides were matched to two *Wolbachia* proteins from our comparative database (see MS proteins 1591 and 1637, **Table 6** and **Table S2**). A 23aa peptide found in two separate charge states mapped to a lipoprotein releasing system transmembrane protein, LolC, that is present in several insect *Wolbachia* strains. Two peptides, one 17aa and one 9aa, mapped to an HlyD family secretion protein from the *Wolbachia* endosymbiont of *Culex quinquefasciatus*. Thus far, we have not identified any genome or transcript sequence capable of producing these peptides in *O. flexuosa*.

Confirmation of Wolbachia-like proteins using peptide antibodies

Polyclonal antibodies raised against the a peptide from MS protein 1591, the putative LolC homolog, detect a single band at 38 kDa in *O. flexuosa* adult worm lysate by Western blot (**Fig. 2**). Total IgG purified from pre-immune serum or an antibody against the keyhole limpit hemocyanin (KLH) carrier protein did not react with any proteins in this size range. The same antibodies were used in immunohistological studies to localize this protein in *O. flexuosa* worms (**Fig. 3**). Similar to the pre-immune total IgG (not shown), antibodies against KLH showed no specific staining in O. flexuosa (**Fig. 3A**). In contrast, the LolC antibody produced intense staining in fibrillar portions of the somatic muscles of adult male (**Fig. 3B**) and young female (**Fig. 3C, D**) worms. Older females have less pronounced somatic musculature and show weaker labeling; however, staining is seen in the uterine muscles (**Fig. 3C, D, F**) and in coiled and stretched microfilaria (**Fig. 3F, G**) within the uteri. Distinct staining is also seen in the membrane of the

excretory cell (**Fig. 3D, E**). Sequences homologous to LolC were identified from the genome of the *Wolbachia* endosymbiont of *B. malayi* and used to make RNA probes for in situ hybridization since corresponding sequences have not yet been identified from *O. flexuosa*. *In situ* hybridization studies indicate that the LolC RNA may be produced in the lateral chords (**Fig. 3I, J, K**), the hypodermis (**Fig. 3K**), developing sperm (**Fig. 3I**), uteri (**Fig. 3K**) and intestine (**Fig. 3K**).

DISCUSSION

Research on filarial nematodes has predominantly focused on *Wolbachia*-dependent species that parasitize humans or domestic animals. *Wolbachia*-free species have historically been overlooked since most are animal parasites without economic or medical importance. However, genetic characterization of *Wolbachia*-free species will lead to a better understanding of filarial biology and may provide insight into the reasons that some species are able to survive in the absence of *Wolbachia* while others cannot.

In this study, we chose to focus on the transcriptome rather than the genome of *O. flexuosa* as a cost-effective method for exploring the genetic complement of a *Wolbachia*-free filarial parasite. The production of 16,814 isogroups and 47,252 high-quality singletons represents a significant increase in the data available for this species, as only 2,124 ESTs were previously available. Our estimates of completeness suggest that this dataset may represent as many as 41% of the genes encoded by the *O. flexuosa* genome. For comparison, 93% of CEGs were identified from *Ancylostoma caninum* transcripts in a recent study, but only after sequencing four different life cycle stages [21]. *O. flexuosa* nodules generally contain one or a few large adult female worms and their developing embryos along with the occasional male [16]. While it is likely that our cDNA library, derived from the contents of two nodules, contains some diversity (i.e. adult females and males of different ages, embryos and MF in various developmental stages), it is heavily biased towards transcripts expressed in adult females. It will be necessary to analyze male parasites as well as other life cycle stages, particularly those found in the insect vector, in order to discover more genes through transcriptome

sequencing. Considering the fact that *O. flexuosa* is a wild-caught species that cannot be maintained in the lab, this material may be difficult to obtain.

Important biological information can be inferred from transcriptome data; however, it is the proteome, not the transcriptome, which mediates cellular function. Unfortunately, shotgun proteomic analyses are only informative if sequences from the organism of interest are available so that MS peptides can be characterized by comparison. Sequences from closely related organisms can be used when genomic data are unavailable, but only highly conserved sequences will be informative [22,23,24]. Transcriptome data can also be used in lieu of genome sequences [23,25], but transcriptome datasets tend to be incomplete, heavily fragmented, and poorly annotated. Our combined approach, using both orthologous sequences and transcriptome data, resulted in the identification of 1,803 proteins. In this instance, sequence conservation between *O. flexuosa* and other filarial species allowed for a high rate of cross-species matches (~60% of matches). Previous studies have shown that very few peptides will match exactly given an overall sequence identity less than 70% [26,27]; therefore, it is not surprising that relatively few matches were made to *C. elegans* sequences given that *O. flexuosa* and *C. elegans* HSPs shared an average of 70.0% identity. The higher rates of identity shared between *O. flexuosa* and *B. malayi* (80.7%) allowed for a greater number of matches.

In this study, peptide translations and experimentally verified proteins were classified based on the presence of conserved protein domains and by comparison with the KEGG database. The most common functional domains among *O. flexuosa* predicted proteins (collagen repeats, kinase domains and WD40 repeat domains) are also common among the transcriptomes of other parasitic nematodes, including *Trichostrongylus*

colubriformis, *Ostertagia ostertagi* and *Necator americanus* [28,29,30] and the proteome of *B. malayi* [13]. Thioredoxin-like folds, motifs common among proteins that catalyze the formation and/or isomerization of disulfide bonds, were the most commonly identified domains among experimentally verified proteins. Immunohistological studies have shown that thioredoxin peroxidase is abundant in the body wall, uterus and intestine of adult worms and on the surface of MF [31]. Antioxidant proteins like these may be required in abundance to resist reactive oxygen species produced by host granulocytes and macrophages present in *O. flexuosa* nodule tissue [16]. The discrepancies seen domains predicted from the transcriptome and proteome may be a result of differences in expression at the RNA and protein levels or they may be due to variance between the nodules (e.g. age of worms, number of males and females, embryo/MF content, etc) used in the transcriptomic and proteomic studies.

Using KEGG pathway modules, we investigated several biochemical pathways that are present in the *Wolbachia* endosymbiont of *B. malayi* but absent in *B. malayi* [11,12,32], including those related to the de novo synthesis of heme, riboflavin, purines and pyrimidines. *Wolbachia*-dependent filaria may use *Wolbachia* as a source for one or more of these substances. If this were true, a *Wolbachia*-free worm like *O. flexuosa* would need to synthesize these substances directly or obtain them from an alternative source (e.g. the mammalian host instead of the bacterial endosymbiont). While our present dataset provides little evidence to support this hypothesis with respect to heme or riboflavin, it is interesting that the partial transcriptome of *O. flexuosa* contains several sequences related to nucleotide synthesis (**Table 5**). Future studies will be required to

determine if *O. flexuosa* encodes complete and functional pathways for the synthesis of purines and pyrimidines.

In addition to providing a general survey of the transcriptome and proteome of *O. flexuosa*, another major goal of this project was to look for expressed *Wolbachia*-like genes or gene fragments. Indeed, many *Wolbachia*-like sequences were found among the *O. flexuosa* transcripts. Only a few of these were identified in a prior survey of the *O. flexuosa* genome even though 56 of the *Wolbachia*-like sequences found in the genome appeared to be expressed based on qRT-PCR studies [15]. This disparity may be due to variation in the wild caught samples used in the two studies, differences in detection sensitivity between qRT-PCR and 454 sequencing, or the use of poly(A) selected RNA in the present study versus total RNA in the previous study.

Regardless of the exact *Wolbachia* genes represented, many similarities exist between the *Wolbachia*-like sequences identified in the genome and transcriptome surveys (see [15] and **Table S3**). Sequences from both studies share homology with genes from both insect and filarial *Wolbachia* strains. While some of the sequences fall within computationally predicted open reading frames, others are disrupted by stop codons and frameshift mutations despite being expressed at the RNA level. As in the previous study [15], we detect no overall theme in the predicted functions of *Wolbachia*-like sequences based on comparison with the KEGG database. It is likely that biochemical themes will not emerge until a draft of the *O. flexuosa* has been generated. However, it is reasonable to assume that *Wolbachia*-like sequences that have been maintained over the course of time and are expressed as RNAs and/or proteins may have an important functions.

We have not yet identified a full-length, uninterrupted *Wolbachia*-like gene in *O. flexuosa*. This may be due to technical issues (e.g., insufficient coverage, fragmentation, etc.). However, finding *Wolbachia*-like and nematode like sequences on the same transcripts suggests that the *Wolbachia*-like sequences may serve as functional domains that alter or enhance the function of existing nematode genes. Many proteins are made up of multiple domains whose autonomous functions contribute to the overall function of the protein as a whole, and recent studies have suggested that domains rather than whole proteins might be common units of horizontal transfer [33].

Our prediction that *Wolbachia*-like sequences play a role in the biology of *O. flexuosa* was strengthened by the finding that two *Wolbachia*-like sequences were expressed at the protein level. Putative homologs of *Wolbachia* HlyD and LolC were identified in adult worm lysate by MS, and the expression of the putative LolC homolog was verified using polyclonal antibodies against the MS peptide. Neither of these proteins has been studied in *Wolbachia*, but both are related to ABC transport in *Escherichia coli* and other bacteria. HlyD is a transenvelope component of the hemolysin secretion system responsible for carrying alpha-hemolysin through the periplasmic space [34]. LolC is an integral membrane component of the LolCDE complex, a transporter that shuttles lipoproteins to the outer membrane [35]. ABC transporters are common to both prokaryotes and eukaryotes. Thirty-three ABC transport proteins have been identified in *B. malayi* [36]. They interact with a diverse array of substrates, and they may play a role in drug resistance in filarial nematodes [37,38,39,40]. Regardless, the exact functions of HlyD and LolC in *O. flexuosa* remain to be seen.

We hypothesize that technical issues may have limited our ability to detect more *Wolbachia*-like proteins. Two main issues hindered our proteomic search: the suboptimal comparative database (discussed previously) and the complexity of our analyzed sample. Even in infected worms, *Wolbachia* proteins can be difficult to detect given that they represent such a small fraction of a worm's total body mass. Despite the thorough analysis of several life cycle stages of *B. malayi*, Bennuru et al. only detected ~70% of the proteins encoded by its *Wolbachia* endosymbiont, and many of these identifications were based on single peptides [13]. Detection of *Wolbachia* proteins from both infected and uninfected filarial species might benefit from a more directed approach, as certain tissues (i.e. lateral chords, ovaries, developing sperm) or life cycle stages (i.e. adolescent worms) may express these genes to a higher degree than others (see [15,41] and **Fig. 3**). Analyses of isolated tissues might lead to a higher identification rates.

Our immunohistological and *in situ* hybridization studies revealed interesting, tissue-specific patterns of expression. To date, three *Wolbachia* like transcripts (homologs of 2-methylthioadenine synthase and DNA polymerase I described in [15] and LolC shown in **Fig. 3**) have been localized, and all three exhibit very similar expression patterns. Probes against the corresponding RNAs stained tissues that would typically harbor *Wolbachia* in an infected species [9,41]. However, the putative LolC homolog was mainly localized to fibrillar portions of the muscles. mRNAs and their corresponding proteins are generally found in the same tissue, but there are precedents for disparate localization. A classic example is a microfilarial sheath protein that is expressed in the uterine epithelium and deposited onto developing microfilaria [42,43]. It is possible that proteins manufactured by *Wolbachia* within infected tissues are exported for use in other locations in the

Wolbachia-dependent worm. Our results suggest that tissues known to harbor *Wolbachia* in infected worms (i.e., lateral chords and reproductive organs) may also be responsible for the production of *Wolbachia*-like products in *Wolbachia*-free worms, despite the transport of these products to other tissues. This special function of lateral chord and reproductive tissue may be vital in the worm's ability to molt and reproduce, respectively.

This study has provided new information on the transcriptome and proteome of *O. flexuosa* and additional evidence of the expression of *Wolbachia*-like genes in an endosymbiont-free filarial species. Our proteomic analysis also provided the first proof that horizontally transferred *Wolbachia* genes are expressed at the protein level. Additional studies will be needed to determine the role(s) of *Wolbachia*-like proteins in *Wolbachia*-free filarial nematodes.

MATERIALS AND METHODS

Parasite material

Adult *O. flexuosa* worms were collected from European red deer (*Cervus elaphus*) in northern Germany (Schleswig-Holstein) as previously described [16]. Nodules were removed from the skins of freshly shot deer and frozen at -80°C for use in proteomic studies or fixed in 4% buffered formalin, embedded in paraffin and sectioned according to standard histological technique for use in *in situ* hybridization or immunohistology. Worms used for RNA extraction were dissected from fresh nodules and stored in TRIzol (Invitogen, Carlsbad, CA, USA) at -80°C.

Adult *B. malayi* were obtained from experimentally infected Mongolian gerbils as previously described [44]. Worms were stored at -80°C prior to use.

RNA Isolation, library construction and 454 sequencing

Total RNA was isolated from *O. flexuosa* worm fragments, DNase treated and tested for DNA contamination as previously described [15]. RNA yield and integrity were assessed using a NanoDrop ND-1000 UV-VIS spectrophotometer (NanoDrop Technologies, Wilmington, DE) and a Bioanalyzer 2100 (Agilent Technologies, Cedar Creek, TX), respectively. Full-length cDNA was generated from 1.0µg total RNA using an optimized 27-cycle protocol with the Accuscript HF Reverse Transcriptase Kit (Agilent) and SMART primers (Invitrogen). The resulting cDNA library was normalized with the Trimmer kit (Evrogen, San Diego, CA) and amplified over 14 cycles using SMART primers (Invitrogen) and Clontech Advantage-HF 2 polymerase (Clontech/Takara Bio, CA). 300-800bp fragments were selected using AMPure

paramagnetic beads (Agencourt, Beckman Coulter Genomics, Beverly, MA) following removal of 3' and 5' adaptors by restriction digest and processed using the Titanium General Library Kit (Roche, Branford, CT). Sequencing was performed on the Genome Sequencer 454 Titanium instrument using the GS FLX Titanium Sequencing Kit (Roche) according to standard protocol [45].

Assembly and translation of transcript sequences

Raw 454 reads were trimmed for adapters and low complexity regions using SeqClean (http://www.tigr.org/tdb/tgi/software) and screened against UniVec (http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html) and genome sequences from *Bos taurus* (the closest sequenced relative of the deer host) and *Homo sapiens*. The Newbler assembler [45] was used to assemble clean reads and 2,124 Genbank ESTs (September 2010) using the following parameters "-cdna –ml 100 –mi 95 –icl –het". Any isotig or singleton less than 200bp, with more than 10% ambiguous bases, or with a top BLAST hit to a human or Ruminant nucleotide or protein sequence was excluded from analysis. Peptide translations were obtained from isotigs and singletons using Prot4EST [46,47].

Core eukaryotic gene mapping and assessment of fragmentation

The coverage of the *O. flexuosa* transcriptome was estimated by profile searching isotigs and singletons against core eukaryotic genes [18,19] using HMMer [48]. A custom perl script was used to calculate the fragmentation rate based on WU-BLAST (http://blast.wustl.edu) BLASTX alignments to *B. malayi* (brugpep.WS225.fa) and *C. elegans* (wormpep.WS21.fa) proteins.

BLAST analyses

All BLAST searches were performed using blastall version 2.2.22 with the following cutoff values: e-values less than 1e-05, bit scores greater than 35 bits, and percent identity greater than 55%. Databases queried included: non-redundant protein and non-redundant nucleotide databases (downloaded 5/27/2010), all Genbank ESTs related to the Nematoda (downloaded 1/6/2011), WormBase *B. malayi* genome assembly (b_malayi.WS221.dna.fa), WormBase *B. malayi* predicted proteins (brugpep.WS221.fa), WormBase *C. elegans* predicted proteins (wormpep.WS215.fa), and the KEGG protein database (version 57). A custom perl script was used to extract the top hit from each region of the query before parsing BLAST results.

Mass spectroscopy

O. flexuosa worms were dissected from a frozen nodule, pulverized in liquid nitrogen, dissolved in 50mM Tris pH 8.0 with cOmplete protease inhibitor cocktail (Roche), and sonicated on ice. Two nodules were prepared in this manner to provide biological replicates. Half of each sonicated sample was pooled and subjected to centrifugation at 20,000 x g for 90 minutes to separate soluble from insoluble proteins. This resulted in four separate samples (nodule 1, nodule 2, nodule 1-2 soluble fraction, and nodule 1-2 insoluble fraction) that were treated in the same manner. Protein concentration was determined by the DC Protein Assay (Bio Rad, Hercules, CA). Proteins were denatured with 8M urea, reduced with 5 mM tris (2-carboxyethyl) phosphine (TCEP), and alkylated with iodoacetamide, prior to digestion with sequencing-grade trypsin (1:50 trypsin:protein ratio).

Peptides from the four samples were separated using multidimensional chromatography (MudPIT) and analyzed by tandem MS as previously described [49]. Briefly, 50 μg of peptides were pressure-loaded onto a biphasic capillary column packed with strong cation exchange (SCX) and C18 resins, and fractionated by a combination of ion exchange and reverse phase chromatography. Peptides were eluted from SCX resin with six injections of increasing concentrations of ammonium acetate, followed by 120 minute organic gradients to elute peptides from the C18 resin directly into an LTQ Orbitrap Discovery hybrid mass spectrometer (Thermo Scientific, West Palm Beach, FL). MS1 spectra were acquired in the orbitrap (FTMS) with a resolution of 30,000 followed by seven data-dependent MS2 spectra in the ion trap (ITMS) at low resolution. Data were stored in Thermo RAW format, and converted to MS2 format using RawXtract 1.9.9.2 [50].

MS2 spectra were searched using the ProLuCID algorithm [51] against a combined database containing 238,403 sequences derived from *O. flexuosa* peptide translations, all GenBank protein entries from the phylum Nematoda and genus *Wolbachia* (downloaded February 2010), *WormBase* protein sequences from *B. malayi* (brugpep.WS221.fa) and *C. elegans* (wormpep.WS215.fa). Reversed "decoy" versions of each entry were also included to allow for estimation of false-discovery rates [52]. Peptide spectral matches were filtered using DTASelect 2.0 [53] resulting in a peptide false-positive rate of 1.2%.

Functional annotation of predicted and experimentally verified proteins

InterProScan version 4.5 was used to identify conserved domains from peptide translations and database protein sequences with matches to MS peptides [54]. Sequences

were assigned K numbers by comparison with the KEGG protein database [55]. Custom perl scripts were used to bin K numbers into broad categories and KEGG modules.

Antibody production

A 15 amino acid portion of MS protein 1591 (see **Table 6**) was selected based on inferred immunogenic properties (i.e. predicted secondary structure and hydrophobicity). Polyclonal antisera were raised against synthetic peptides coupled to KLH in two rabbits per peptide (LifeTein LLC, South Plainfield, NJ). Antibodies were affinity purified with the original peptide and tested by ELISA prior to use (LifeTein LLC). Polyclonal antisera against the KHL carrier protein were raised, purified, and tested in the same manner. Total IgG were purified from rabbit pre-immune sera using the Protein A Agarose Kit (KPL, Gaithersburg, MD) for use as negative controls.

Western blots

Adult *O. flexuosa* were dissected from nodules and homogenized on ice in RIPA buffer (G Biosciences, Maryland Heights, MO) and ProteaseArrest (G Biosciences) in a 1 mL mini homogenizer (GPE Scientific Limited, Leighton Buzzard, UK). Homogenate was spun at 19,000g for 15 minutes to pellet debris, and the protein concentration of the supernatant was determined using the BCA method (Pierce, Rockford, IL). *O. flexuosa* protein was subjected to SDS-PAGE using a 4-12% reducing gel (NuPAGE BisTris Mini Gel, Invitrogen) according to the manufacturer's suggested protocol. Separated proteins were transferred to a nitrocellulose membrane (Invitrogen). The membrane was blocked overnight at 4C in blocking buffer (0.5% Tween, 5% nonfat dry milk in 1xPBS), washed

in 1xPBS with 0.5% Tween (PBS/T), and cut into strips for blotting. Antibodies were diluted in PBS/T to final concentrations of 4.4 μg/mL,15 μg/mL and 50 μg/mL for anti-MS1591, purified pre-immune IgG and anti-KLH, respectively. Blot strips were incubated with primary antibody dilutions overnight at 4C and washed with PBS/T at room temperature. Strips were then incubated with anti-rabbit IgG(Fc) AP conjugate (1:3,500 in PBS/T) (Promega, Sunnyvale, CA) for 1h at 37C, washed with PBS/T at room temperature and developed using NBT/BCIP substrate (Promega). Substrate reaction was stopped using 20mM Tris-HCl pH 7.4 with 5mM EDTA.

In situ hybridization and immunohistochemistry

DNA was isolated from adult *B. malayi* as previously described [15] using the DNeasy Blood and Tissue Kit (Qiagen). A 369bp portion of the LolC gene from the *Wolbachia* endosymbiont of *B. malayi* (locus Wbm0482) was amplified from *B. malayi* genomic DNA using the following primers: 5'- TCTTTCATTCTCGGCACCTCA-3' and 5'- TGGCATTGATGGCCATATCA-3'. Biotinylated RNA probes were constructed from PCR products by in vitro transcription as previously described [15].

O. flexuosa nodules were fixed for 24-72 hours in DEPC-treated 4% buffered formaldehyde, embedded in paraffin and sectioned using standard histological procedures. Immunohistological stainings and *in situ* hybridizations were performed on sections from the nodules as previously described [41]. *In situ* hybridizations were performed overnight at 60C using 1μg RNA probe. Antibodies against MS protein 1591 (LolC) used at 0.88 μg/mL while anti-KLH antibodies were used at 1.4 μg/mL. A total of twelve nodules were examined.

ACKNOWLEDGEMENTS

We would like to thank Jeffrey I. Gordon (Center for Genome Sciences and Systems Biology, Washington University) for access to the mass spectrometer and related reagents and Katherine Mann for technical assistance. This work was supported in part by a grant from the Barnes-Jewish Hospital Foundation.

REFERENCES

1. WHO (2009) Global programme to eliminate lymphatic filariasis. Wkly Epidemiol Rec 84: 437-444.
2. WHO (2010) African Programme for Onchocerciasis Control - report of the sixth meeting of national task forces, October 2009. Wkly Epidemiol Rec 85: 23-28.
3. Hoerauf A, Mand S, Fischer K, Kruppa T, Marfo-Debrekyei Y, et al. (2003) Doxycycline as a novel strategy against bancroftian filariasis-depletion of *Wolbachia* endosymbionts from *Wuchereria bancrofti* and stop of microfilaria production. Med Microbiol Immunol 192: 211-216.
4. Hoerauf A, Nissen-Pahle K, Schmetz C, Henkle-Duhrsen K, Blaxter ML, et al. (1999) Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility. J Clin Invest 103: 11-18.
5. Slatko BE, Taylor MJ, Foster JM (2010) The *Wolbachia* endosymbiont as an anti-filarial nematode target. Symbiosis 51: 55-65.
6. Hoerauf A, Mand S, Volkmann L, Buttner M, Marfo-Debrekyei Y, et al. (2003) Doxycycline in the treatment of human onchocerciasis: Kinetics of *Wolbachia* endobacteria reduction and of inhibition of embryogenesis in female *Onchocerca* worms. Microbes Infect 5: 261-273.
7. Supali T, Djuardi Y, Pfarr KM, Wibowo H, Taylor MJ, et al. (2008) Doxycycline treatment of *Brugia malayi*-infected persons reduces microfilaremia and adverse reactions after diethylcarbamazine and albendazole treatment. Clin Infect Dis 46: 1385-1393.
8. Turner JD, Tendongfor N, Esum M, Johnston KL, Langley RS, et al. (2010) Macrofilaricidal activity after doxycycline only treatment of *Onchocerca volvulus* in an area of *Loa loa* co-endemicity: a randomized controlled trial. PLoS Negl Trop Dis 4: e660.
9. Ferri E, Bain O, Barbuto M, Martin C, Lo N, et al. (2011) New Insights into the Evolution of *Wolbachia* Infections in Filarial Nematodes Inferred from a Large Range of Screened Species. PLoS One 6: e20843.
10. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. Int J Parasitol 34: 191-203.
11. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, et al. (2005) The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. PLoS Biol 3: e121.
12. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al. (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317: 1756-1760.
13. Bennuru S, Meng Z, Ribeiro JM, Semnani RT, Ghedin E, et al. (2011) Stage-specific proteomic expression patterns of the human filarial parasite *Brugia malayi* and its endosymbiont *Wolbachia*. Proc Natl Acad Sci U S A 108: 9649-9654.
14. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753-1756.

15. McNulty SN, Foster JM, Mitreva M, Dunning Hotopp JC, Martin J, et al. (2010) Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. PLoS One 5: e11029.
16. Plenge-Bonig A, Kromer M, Buttner DW (1995) Light and electron microscopy studies on *Onchocerca jakutensis* and *O. flexuosa* of red deer show different host-parasite interactions. Parasitol Res 81: 66-73.
17. Martin J, Abubucker S, Heizer E, Taylor CM, Mitreva M (2012) Nematode.net update 2011: addition of data sets and tools featuring next-generation sequencing data. Nucleic Acids Res 40: D720-728.
18. Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23: 1061-1067.
19. Parra G, Bradnam K, Ning Z, Keane T, Korf I (2009) Assessing the gene space in draft genomes. Nucleic Acids Res 37: 289-297.
20. Ramani AK, Calarco JA, Pan Q, Mavandadi S, Wang Y, et al. (2011) Genome-wide analysis of alternative splicing in *Caenorhabditis elegans*. Genome Res.
21. Wang Z, Abubucker S, Martin J, Wilson RK, Hawdon J, et al. (2010) Characterizing *Ancylostoma caninum* transcriptome and exploring nematode parasitic adaptation. BMC Genomics 11: 307.
22. Carpentier SC, Panis B, Vertommen A, Swennen R, Sergeant K, et al. (2008) Proteome analysis of non-model plants: a challenging but powerful approach. Mass Spectrom Rev 27: 354-377.
23. Moreno Y, Gros PP, Tam M, Segura M, Valanparambil R, et al. (2011) Proteomic analysis of excretory-secretory products of *Heligmosomoides polygyrus* assessed with next-generation sequencing transcriptomic information. PLoS Negl Trop Dis 5: e1370.
24. Buggiotti L, Primmer CR, Kouvonen P, Bures S, Corthals GL, et al. (2008) Identification of differentially expressed proteins in *Ficedula* flycatchers. Proteomics 8: 2150-2153.
25. Wang H, Zhang H, Wong YH, Voolstra C, Ravasi T, et al. (2010) Rapid transcriptome and proteome profiling of a non-model marine invertebrate, *Bugula neritina*. Proteomics 10: 2972-2981.
26. Lester PJ, Hubbard SJ (2002) Comparative bioinformatic analysis of complete proteomes and protein parameters for cross-species identification in proteomics. Proteomics 2: 1392-1405.
27. Wilkins MR, Williams KL (1997) Cross-species protein identification using amino acid composition, peptide mass fingerprinting, isoelectric point and molecular mass: a theoretical evaluation. J Theor Biol 186: 7-15.
28. Abubucker S, Zarlenga DS, Martin J, Yin Y, Wang Z, et al. (2009) The transcriptomes of the cattle parasitic nematode *Ostertagia ostartagi*. Vet Parasitol 162: 89-99.
29. Cantacessi C, Mitreva M, Campbell BE, Hall RS, Young ND, et al. (2010) First transcriptomic analysis of the economically important parasitic nematode, *Trichostrongylus colubriformis*, using a next-generation sequencing approach. Infect Genet Evol 10: 1199-1207.

30. Cantacessi C, Mitreva M, Jex AR, Young ND, Campbell BE, et al. (2010) Massively parallel sequencing and analysis of the *Necator americanus* transcriptome. PLoS Negl Trop Dis 4: e684.
31. Lu W, Egerton GL, Bianco AE, Williams SA (1998) Thioredoxin peroxidase from *Onchocerca volvulus*: a major hydrogen peroxide detoxifying enzyme in filarial parasites. Mol Biochem Parasitol 91: 221-235.
32. Scott AL, Ghedin E (2009) The genome of *Brugia malayi* - all worms are not created equal. Parasitol Int 58: 6-11.
33. Chan CX, Darling AE, Beiko RG, Ragan MA (2009) Are protein domains modules of lateral genetic transfer? PLoS One 4: e4524.
34. Fath MJ, Kolter R (1993) ABC transporters: bacterial exporters. Microbiol Rev 57: 995-1017.
35. Yakushi T, Masuda K, Narita S, Matsuyama S, Tokuda H (2000) A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. Nat Cell Biol 2: 212-218.
36. Ardelli BF, Stitt LE, Tompkins JB (2010) Inventory and analysis of ATP-binding cassette (ABC) systems in *Brugia malayi*. Parasitology 137: 1195-1212.
37. Ardelli BF, Prichard RK (2004) Identification of variant ABC-transporter genes among *Onchocerca volvulus* collected from ivermectin-treated and untreated patients in Ghana, West Africa. Ann Trop Med Parasitol 98: 371-384.
38. Ardelli BF, Prichard RK (2007) Reduced genetic variation of an *Onchocerca volvulus* ABC transporter gene following treatment with ivermectin. Trans R Soc Trop Med Hyg 101: 1223-1232.
39. Bourguinat C, Ardelli BF, Pion SD, Kamgno J, Gardon J, et al. (2008) P-glycoprotein-like protein, a possible genetic marker for ivermectin resistance selection in *Onchocerca volvulus*. Mol Biochem Parasitol 158: 101-111.
40. Tompkins JB, Stitt LE, Morrissette AM, Ardelli BF (2011) The role of *Brugia malayi* ATP-binding cassette (ABC) transporters in potentiating drug sensitivity. Parasitol Res 109: 1311-1322.
41. Fischer K, Beatty WL, Jiang D, Weil GJ, Fischer PU (2011) Tissue and Stage-Specific Distribution of *Wolbachia* in *Brugia malayi*. PLoS Negl Trop Dis 5: e1174.
42. Jiang D, Li BW, Fischer PU, Weil GJ (2008) Localization of gender-regulated gene expression in the filarial nematode *Brugia malayi*. Int J Parasitol 38: 503-512.
43. Dafa'alla TH, Taubert A, Hobom G, Beck E, Zahner H (2000) Molecular characterization of a *Litomosoides sigmodontis* protein involved in the development of the microfilarial sheath during embryogenesis. Mol Biochem Parasitol 106: 37-50.
44. Ash LR, Riley JM (1970) Development of subperiodic *Brugia malayi* in the jird, *Meriones unguiculatus*, with notes on infections in other rodents. J Parasitol 56: 969-973.
45. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376-380.
46. Wasmuth J, Blaxter M (2009) Obtaining accurate translations from expressed sequence tags. Methods Mol Biol 533: 221-239.

47. Wasmuth JD, Blaxter ML (2004) prot4EST: translating expressed sequence tags from neglected genomes. BMC Bioinformatics 5: 187.
48. Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14: 755-763.
49. Washburn MP, Wolters D, Yates JR, 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19: 242-247.
50. McDonald WH, Tabb DL, Sadygov RG, MacCoss MJ, Venable J, et al. (2004) MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. Rapid Commun Mass Spectrom 18: 2162-2168.
51. Xu T, Venable JD, Park SK, Cociorva D, Lu B, et al. (2006) ProLuCID, a Fast and Sensitive Tandem Mass Spectra-based Protein Identification Program. Mol Cell Proteomics 5: S174.
52. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 4: 207-214.
53. Tabb DL, McDonald WH, Yates JR, 3rd (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res 1: 21-26.
54. Zdobnov EM, Apweiler R (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17: 847-848.
55. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28: 27-30.

FIGURES:

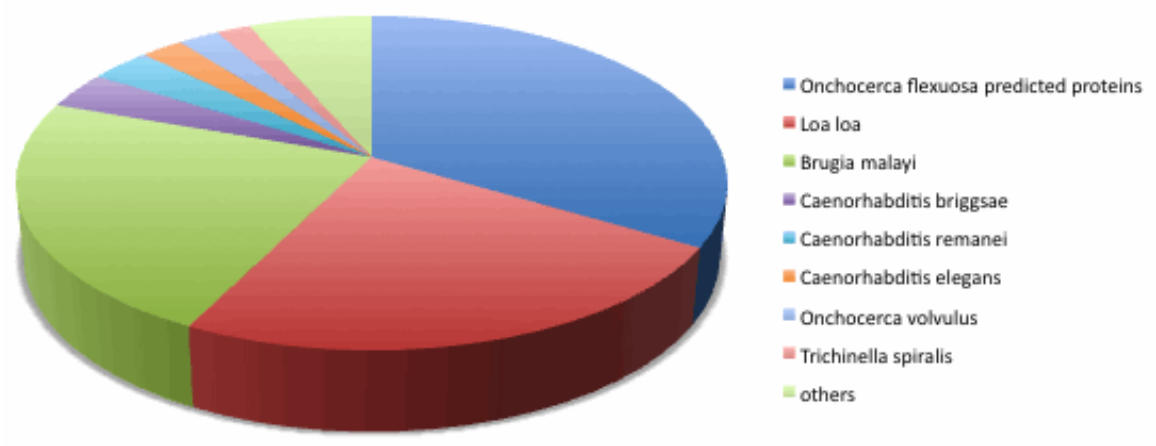


Figure 1: Cross species identification rates of *O. flexuosa* proteins detected by mass spectrometry. Approximately 30% of the experimentally detected proteins mapped to *O. flexuosa* peptide translations. 22% of proteins were matched to *L. loa* while another 22% matched to *B. malayi*. Fewer matches were made to distantly related species (approximately 3% to *C. briggsae*, 3% to *C. remanei*, 3% to *C. elegans*, and 2% to *T. spiralis*) and to closely related species that are not well represented in the GenBank protein database (2% to *O. volvulus*).

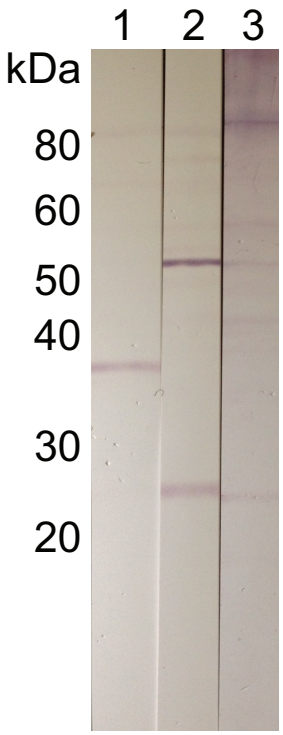


Figure 2: Western blots detecting the putative LolC homolog in *O. flexuosa* adult worm lysate. Affinity-purified rabbit antibodies raised against a peptide from MS protein 1591, a putative LolC homolog, bound a single 38 kDa band in Western blots of *O. flexuosa* adult worm lysate (lane 1). Purified IgG from pre-immune serum (lane 2) and antibodies against the keyhole limpit hemocyanin carrier (lane 3) did not bind to proteins in this size range.

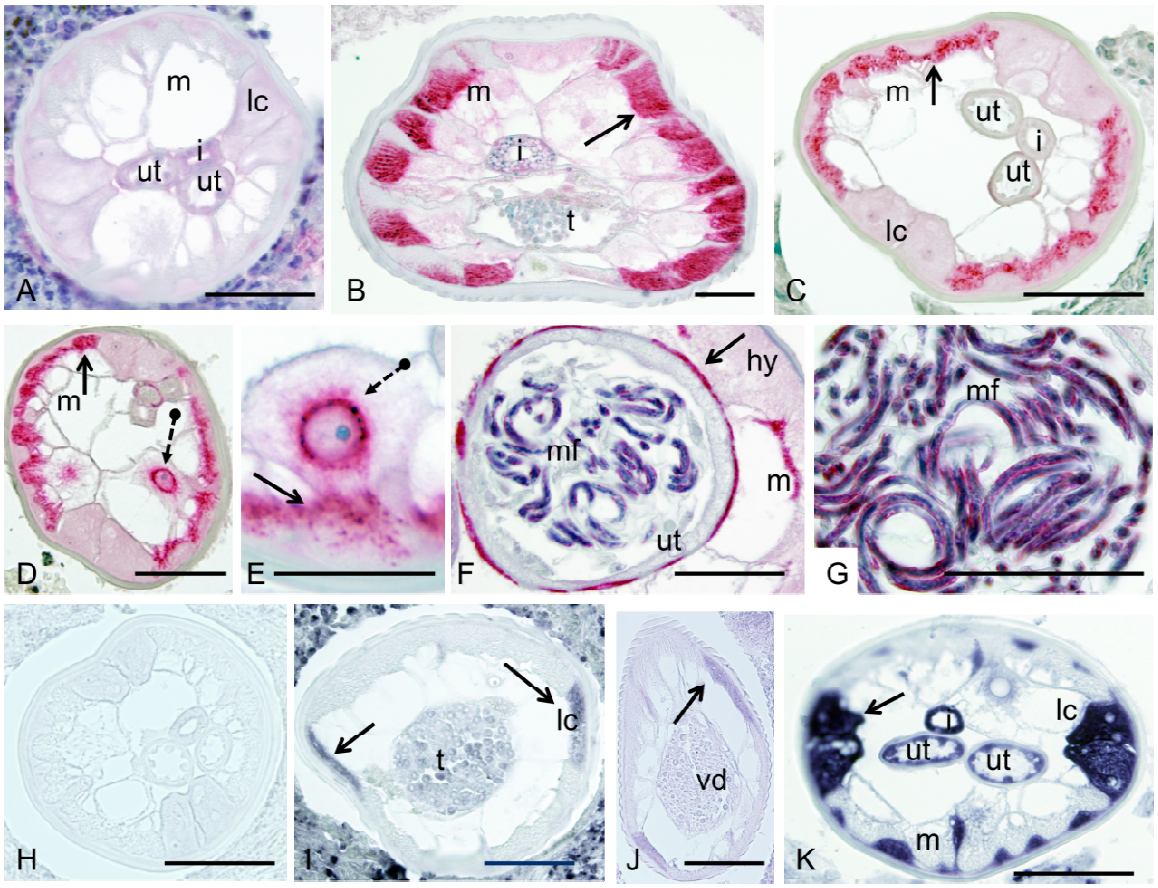


Figure 3: Localization of the putative LolC homolog by immunohistology and *in situ* hybridization in *O. flexuosa*. Negative control antibodies against keyhole limpit hemocyanin (carrier protein) showed no specific staining in *O. flexuosa* (**A**). Antibodies against a peptide from MS protein 1591, a putative homolog of *Wolbachia* LolC, stained the fibrillar portions of the somatic muscles (arrow) in a cross section of an adult male worm (**B**). The somatic muscles (solid arrows) and excretory cell (dashed arrow) are stained in cross sections of young adult females (**C, D**). Magnification of the excretory cell shows intense staining of the cell membrane (dashed arrow) and more diffuse staining in adjacent muscles (solid arrow) (**E**). The uterine muscles of an older female (arrow) are clearly stained (**F**), as are the intrauterine stretched microfilaria (**F, G**). The LolC sense probe produced no signal in *in situ* hybridizations (**H**), while the antisense probe labeled the lateral chords (arrows) and developing sperm within the male testes (**I, J**). The antisense probe also labeled the lateral chords (arrow), intestine, and uteri of a young adult female (**K**). Abbreviations: m, muscle; lc, lateral chords; i, intestine; ut, uterus; mf, microfilariae; hy, hypodermis; t, testes; vd, vas deferens. Scale bars = 25µm.

TABLES:

Table 1: Results of sequencing, assembly, and translation of the *O. flexuosa* transcriptome.

Sequencing	
Total reads	1,334,767
Average read length	350 bp
Total ESTs (Genbank)	2,124
Average EST length	443 bp
Assembly	
# Singletons	47,252
Average singleton length	538 bp
# Isotigs	25,222
Average isotig length	810 bp
# Isogroups	16,814
# Isogroups with one isotig	12,874
Average isotigs per AS isogroup	3.13
Protein Predictions	
# Unique peptides predicted	68,402
# Isogroups with predicted peptides	16,807
# Isotigs with predicted peptides	25,205
# Singletons with predicted peptides	47,252

Table 2: Results of BLAST searches comparing *O. flexuosa* transcripts to various databases.

Database	Database Source	Sequence search	# Isogroups	# Isotigs	# Singletons
nr	NCBI	blastx	6,619 (39.4%)	12,013 (47.6%)	7,278 (15.4%)
Nematoda			6,442	11,755	6,886
Wolbachia			62	73	36
Nematoda and *Wolbachia*			17	23	3
nt	NCBI	blastn	6,998 (41.6%)	12,641 (50.1%)	9,988 (21.1%)
Nematoda			6,613	12,141	8,731
Wolbachia			86	98	102
Nematoda and *Wolbachia*			20	22	10
Nematoda ESTs	GenBank	blastn	5,485 (32.6%)	9,765 (38.7%)	7,535 (15.9%)
B. malayi proteins	WormBase	blastx	6,683 (39.7%)	12,153 (48.2%)	7,702 (16.3%)
B. malayi genome assembly	WormBase	blastn	8,350 (49.7%)	14,522 (57.6%)	11,836 (25.0%)
C. elegans proteins	WormBase	blastx	2,437 (14.5%)	4,324 (17.1%)	2,385 (5.0%)
Hit to one or more databases			10,024 (59.6%)	16,944 (67.2%)	16,359 (34.6%)

Table 3: Top 25 InterPro domains identified from *O. flexuosa* peptide translations and experimentally verified proteins

Peptide Translations			Experimentally Verified Proteins		
InterPro Code	Description	Isogroups and Singletons	InterPro Code	Description	Proteins
IPR008160	Collagen triple helix repeat	189	IPR012336	Thioredoxin-like fold	43
IPR011009	Protein kinase-like domain	167	IPR008985	Concanavalin A-like lectin/glucanase	41
IPR017442	Serine/threonine-protein kinase-like domain	118	IPR013098	Immunoglobulin I-set	40
IPR000719	Protein kinase, catalytic domain	102	IPR016024	Armadillo-type fold	39
IPR011046	WD40 repeat-like-containing domain	67	IPR016040	NAD(P)-binding domain	35
IPR019781	WD40 repeat, subgroup	66	IPR013783	Immunoglobulin-like fold	33
IPR007087	Zinc finger, C2H2-type	63	IPR002928	Myosin tail	31
IPR006201	Neurotransmitter-gated ion-channel	61	IPR004000	Actin-like	31
IPR017986	WD40-repeat-containing domain	60	IPR001664	Intermediate filament protein	30
IPR015943	WD40/YVTN repeat-like-containing domain	59	IPR012335	Thioredoxin fold	30
IPR002290	Serine/threonine-protein kinase domain	53	IPR001023	Heat shock protein Hsp70	29
IPR001680	WD40 repeat	52	IPR013126	Heat shock protein 70	29
IPR001245	Serine-threonine/tyrosine-protein kinase	50	IPR001079	Galectin, carbohydrate recognition domain	28
IPR019782	WD40 repeat 2	50	IPR016044	Filament	27
IPR000504	RNA recognition motif domain	49	IPR000217	Tubulin	26
IPR001757	ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter	48	IPR003008	Tubulin/FtsZ, GTPase domain	26
IPR000242	Protein-tyrosine phosphatase, receptor/non-receptor type	44	IPR001322	Intermediate filament, C-terminal	25

IPR016040	NAD(P)-binding domain	43	IPR001806	Ras GTPase	25
IPR016024	Armadillo-type fold	40	IPR002017	Spectrin repeat	25
IPR013753	Ras	38	IPR005225	Small GTP-binding protein	25
IPR002110	Ankyrin repeat	36	IPR008280	Tubulin/FtsZ, C-terminal	25
IPR013783	Immunoglobulin-like fold	36	IPR013320	Concanavalin A-like lectin/glucanase, subgroup	25
IPR008271	Serine/threonine-protein kinase, active site	34	IPR013753	Ras	25
IPR012677	Nucleotide-binding, alpha-beta plait	34	IPR017936	Thioredoxin-like	25
IPR019775	WD40 repeat, conserved site	34	IPR018316	Tubulin/FtsZ, 2-layer sandwich domain	25

Table 4: KEGG pathway category mappings for *O. flexuosa* peptide translations and experimentally verified proteins.

	Peptide Translations				Experimentally Verified Proteins			
	Isogroups & Singletons		KO Groups		Proteins		KO Groups	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Metabolism	**1939**	**37.58%**	**749**	**36.55%**	**278**	**45.72%**	**195**	**43.72%**
Carbohydrate Metabolism	287	5.56%	106	5.17%	101	16.61%	53	11.88%
Energy Metabolism	304	5.89%	116	5.66%	86	14.14%	56	12.56%
Lipid Metabolism	160	3.10%	70	3.42%	20	3.29%	17	3.81%
Nucleotide Metabolism	176	3.41%	76	3.71%	23	3.78%	15	3.36%
Amino Acid Metabolism	273	5.29%	93	4.54%	55	9.05%	35	7.85%
Metabolism of Other Amino Acids	87	1.69%	34	1.66%	22	3.62%	15	3.36%
Glycan Biosynthesis and Metabolism	184	3.57%	75	3.66%	13	2.14%	10	2.24%
Metabolism of Cofactors and Vitamins	74	1.43%	44	2.15%	12	1.97%	9	2.02%
Metabolism of Terpenoids and Polyketides	46	0.89%	20	0.98%	7	1.15%	4	0.90%
Biosynthesis of Other Secondary Metabolites	20	0.39%	9	0.44%	7	1.15%	4	0.90%
Xenobiotics Biodegradation and Metabolism	59	1.14%	23	1.12%	14	2.30%	9	2.02%
Enzyme Families	631	12.23%	233	11.37%	51	8.39%	39	8.74%
Genetic Information Processing	**2442**	**47.33%**	**1023**	**49.93%**	**233**	**38.32%**	**191**	**42.83%**
Transcription	634	12.29%	283	13.81%	23	3.78%	18	4.04%
Translation	493	9.56%	215	10.49%	91	14.97%	77	17.26%
Folding, Sorting and Degradation	870	16.86%	350	17.08%	114	18.75%	87	19.51%
Replication and Repair	628	12.17%	263	12.84%	25	4.11%	21	4.71%
Environmental Information Processing	**879**	**17.04%**	**320**	**15.62%**	**89**	**14.64%**	**61**	**13.68%**
Membrane Transport	61	1.18%	38	1.85%	6	0.99%	5	1.12%
Signal Transduction	526	10.20%	175	8.54%	54	8.88%	32	7.17%
Signaling Molecules and Interaction	403	7.81%	142	6.93%	43	7.07%	33	7.40%
Cellular Processes	**1119**	**21.69%**	**379**	**18.50%**	**148**	**24.34%**	**93**	**20.85%**

Transport and Catabolism	368	7.13%	139	6.78%	67	11.02%	46	10.31%
Cell Motility	326	6.32%	87	4.25%	64	10.53%	31	6.95%
Cell Growth and Death	298	5.78%	93	4.54%	10	1.64%	8	1.79%
Cell Communication	288	5.58%	90	4.39%	54	8.88%	28	6.28%
Organismal Systems	**765**	**14.83%**	**252**	**12.30%**	**110**	**18.09%**	**74**	**16.59%**
Human Diseases	**675**	**13.08%**	**252**	**12.30%**	**147**	**24.18%**	**92**	**20.63%**
Poorly Characterized	**48**	**0.93%**	**26**	**1.27%**	**2**	**0.33%**	**2**	**0.45%**
Total	5159		2049		608		446	

Table 5: Selected KEGG pathway modules in *B. malayi* and *O. flexuosa*. Table lists the number of module enzymes represented in each dataset.

Module description	Module number	Total enzymes in module	*B. malayi*	*O. flexuosa*		
			Proteins	Peptide translations	Peptide translations with homology to *Wolbachia*	Experimentally verified proteins
Inosine monophosphate biosynthesis	M00048	20	1	4	1	0
Uridine monophosphate biosynthesis	M00051	13	3	5	3	1
Heme biosynthesis	M00121	9	3	3	1	0
Riboflavin biosynthsis	M00125	5	0	0	0	0

Table 6: *Wolbachia*-like proteins in *O. flexuosa*.

MS Protein	Identified Peptides	Match Accessions	Description	Peptide Antigen
1591	RNINSDYYAVSKSIERIPGILKA	225592870 99034569 58699345 58536811 42410767 225677321	lipoprotein releasing system transmembrane protein, LolC	NINSDYYAVSKSIER
1637	KLLSLLATEARLIAIRG QFAEIEGRV	190570554	secretion protein, HlyD family	n/a

See supplementary CD for the following tables:

Table S1: KEGG pathway modules represented by *O. flexuosa* peptide translations and experimentally verified proteins.

Table S2: Detailed description of *O. flexuosa* proteins identified by mass spectroscopy.

Table S3: *Wolbachia*-like sequences in *O. flexuosa*.

Table S4: KEGG pathway modules represented by *Wolbachia*-like peptides in *O. flexuosa*.

CHAPTER 5:

Discussion

This chapter was composed entirely by SNM. Comments from Peter U. Fischer and Gary J. Weil were incorporated into the final version.

SECTION I: OVERVIEW

The goal of this work was to explore the role of *Wolbachia* in the biology of filarial nematodes by attempting to answer the following question: Why is it that some filarial species depend upon *Wolbachia* endobacteria for growth and reproduction while other species do not? We proposed that the differences in *Wolbachia* dependence seen in filarial nematodes could be explained by either a loss-of-function in *Wolbachia*-infected worms, leading to dependence, or a gain-of-function in *Wolbachia*-dependent worms, leading to independence. Either possibility, or a combination of both, could explain the unusual distribution of *Wolbachia* in the filarial lineage.

It is assumed that *Wolbachia* endobacteria were acquired by filarial nematodes through contact with infected insect vectors [1]. The initial invaders may have become fixed in the filarial lineage via reproductive parasitism, as is common in arthropods [2]. The fitness of a vertically transmitted parasite is inextricably linked to that of the host, so selection would have driven this interaction towards mutualism. However, the filarial worm must have lost some function(s) in order to become fully dependent upon the endobacteria to carry out its life cycle. The distribution and phylogenetic affinities of *Wolbachia* suggest that this event took place in an ancestral species prior to the separation of the Onchocercinae and Dirofilariinae and that uninfected species represent secondary losses of the endosymbiont [3].

We reasoned that the mitochondrial genome would be a good place to look for loss-of-function in *Wolbachia*-dependent worms. *Wolbachia* are closely related to the mitochondrial progenitor. If these two organisms overlap in function, *Wolbachia* may be able to compensate for defects in the mitochondria of dependent species. We searched for

signs of altered function by comparing the mitochondrial genomes of *Wolbachia*-dependent and independent worms. Chapter 2 of this thesis describes the finding that the mitochondrial genomes of these two groups are remarkably similar. Sequences provide no evidence of degeneracy in the mitochondrial genomes of *Wolbachia*-dependent filarial nematodes.

Once the worms had become dependent upon *Wolbachia* endobacteria for growth and reproduction, some gain-of-function would be necessary to support the transition back to a solitary existence. These functions could have evolved over the course of time, or they could have been acquired directly from the endosymbiont via horizontal gene transfer (HGT). Many *Wolbachia*-like sequences have been found in the nuclear genomes of infected insects and filarial nematodes [4]. While it is assumed that these sequences are non-functional in *Wolbachia*-dependent species, the transfer of critical genes could render the once vital endosymbiont redundant and expendable.

As a first step in testing this gain-of-function hypothesis, we searched for evidence of HGT in *Wolbachia*-free filarial species. Chapter 3 describes the finding that *Wolbachia*–like DNA sequences are indeed present in the nuclear genomes of two distantly related *Wolbachia*-free species, *Acanthocheilonema viteae* and *Onchocerca flexuosa*. Follow up studies in *O. flexuosa*, described in Chapter 4, assessed the expression of transferred sequences. The existence of *Wolbachia*-like transcripts and a *Wolbachia*-like protein suggests that at least some of these sequences are capable of mediating biological function.

While this thesis project was focused on the interactions between *Wolbachia* and filarial nematodes, other contributions were made to the field of filarial nematode

research. These included detailed studies indicating that a filarial parasite of the common grackle is *Wolbachia*-free (Appendix 1), the evaluation of a whole genome amplification technique for use with filarial DNA (Appendix 2), the generation of sequence data from several species, and the first proteomic analysis of a *Wolbachia*-free filarial species. The various contributions of this thesis (summarized in **Table 1**) will be discussed individually in the following sections.

SECTION II: *WOLBACHIA*-DEPENDENCE DUE TO LOSS-OF-FUNCTION

The mitochondrial genome was a logical place to begin the search for loss-of-function in *Wolbachia*-dependent filarial species. Aside from the reasons previously discussed (e.g., co-inheritance, shared ancestry, etc), filarial mitochondrial genomes (mtDNAs) are relatively small (~13kb), manageable molecules to sequence and analyze. Prior to the initiation of this project, only three filarial mtDNAs were publically available; all of these were from *Wolbachia*-dependent species [5,6,7]. The three mtDNAs were quite similar in sequence and content; however, we hypothesized that a careful comparison of these sequences with those of *Wolbachia*-free species might reveal defects in the mtDNAs of *Wolbachia*-dependent worms, perhaps involving the deletion, truncation, or mutation of mitochondrial genes.

Five mtDNAs were sequenced over the course of this project, one from a *Wolbachia*-dependent species and four from *Wolbachia*-free species. The mtDNA of a sixth species, also *Wolbachia*-free [8], was also made available by another group [9]. As reported in Chapter 2, the nine filarial mtDNAs are very similar in content, arrangement and sequence. No genes are absent or mutated in *Wolbachia*-dependent worms as

compared to their *Wolbachia*-free counterparts. The only notable sequence-level difference was the re-arrangement of five tRNA's in *Chandlerella quiscali* relative to the other species. The relevance of this finding is unknown, as *C. quiscali* was the most divergent species included in the study [10]. It may even belong to a subfamily (Splendidofilariinae) that diverged prior to the introduction of *Wolbachia* in the filarial lineage [8]. We will not be able to determine whether this rearrangement is a consequence of *C. quiscali*'s *Wolbachia* status or its evolutionary distance from the other species until the mtDNAs of more Splendidofilariinae have been sequenced.

It is possible that *Wolbachia*-dependent and independent filarial nematodes could have important differences in mitochondrial function despite the sequence-level similarity of their mtDNAs. Altered expression of mitochondrial genes could have a marked effect on mitochondrial output. Furthermore, many of the proteins involved in oxidative phosphorylation are encoded by the nuclear genome rather than the mitochondria, and some of these genes could be missing or altered in *Wolbachia*-dependent species. It may even be possible to detect differences in the function of isolated mitochondria from *Wolbachia*-dependent and independent worms by directly measuring their rates of oxidative phosphorylation [11,12]. Experiments like these could provide insight into the effects of *Wolbachia* on energy production in filarial worms.

Certainly, our loss-of-function hypothesis does not have to be restricted to mitochondria or to energy production via oxidative phosphorylation. *Wolbachia*-dependent filarial nematodes are best thought of as a composite organism, as neither the worm nor the bacteria can function apart from the other. Taken together, these two organisms are functionally equivalent to a *Wolbachia*-free worm. Therefore, it stands to

reason that some vital function must be missing from the nuclear genome of *Wolbachia*-dependent filarial species but present in *Wolbachia* and *Wolbachia*-independent species. The best way to identify the genes involved in this vital function would be to compare the complete genome sequences of two closely related filarial species with discordant *Wolbachia* status. *Onchocerca jakutensis* (*Wolbachia*-dependent) and *O. flexuosa* (*Wolbachia*-free) are an attractive pair [13,14]. Since they parasitize the same host (European red deer, *Cervus elaphus*), we could be more confident that differences in gene content result from contact or lack of contact with *Wolbachia* rather than the occupation different environmental niches.

SECTION III: *WOLBACHIA*-INDEPENDENCE DUE TO GAIN-OF-FUNCTION

Wolbachia are organisms in transition; they represent an intermediate step between an independent entity and an organelle. An HGT ratchet mechanism has been proposed to explain the ongoing transfer of genetic material from endosymbiont to host during this transition, thereby reducing the endosymbiont genome [15]. As is the case for eukaryotic mitochondria, the transformation of *Wolbachia* from an invader to an obligate mutualist has involved a great deal of genetic rearrangement including the shuttling of bacterial DNA to the nuclear genome of the host. We reasoned that this strategy might have backfired in some filarial species. Instead of streamlining the bacterial genome, the transfer of critical genes may have rendered the bacteria expendable. The first step towards proving this hypothesis, described in Chapter 3, was the identification of *Wolbachia*-like DNA sequences in the nuclear genomes of two *Wolbachia*-free species. The second step, described in Chapter 4, was the identification of *Wolbachia*-like

transcripts and a *Wolbachia*-like protein in a *Wolbachia*-free filarial species, indicating that these sequences are indeed capable of mediating some biological function.

The identification of *Wolbachia*-like sequences in the nuclear genomes of *Wolbachia*-free filarial species is an important scientific finding regardless of their function. Two hypotheses had previously been proposed to explain the unusual distribution of *Wolbachia* among filarial nematodes. Since closely related species may differ in their *Wolbachia*-status, *Wolbachia* either infected each species individually, skipping over the presently *Wolbachia*–free lineages, or they infected an ancestral worm with current *Wolbachia*-free lineages resulting from secondary loss. Scientists favored the latter explanation in the case of the Onchocercinae and Dirofilariinae because some 90% of these species are infected with *Wolbachia* [3,16] and because the phylogeny of *Wolbachia* closely tracks the phylogeny of their filarial hosts within these two groups [16]. Our discovery of *Wolbachia*-like sequences in the nuclear genomes of *A. viteae* and *O. flexuosa* provided the first concrete evidence to support this hypothesis. The identification of shared *Wolbachia*-like sequences in worms of these two subfamilies would further support the conclusion that the infection occurred in a common ancestor.

The transfer of *Wolbachia* DNA to the nuclear genome of the host was clearly a random process. Many of the *Wolbachia*-like inserts are unexpressed sequence fragments that have no clear importance (see Chapter 3). However, our experimental results provide strong circumstantial evidence that some of the *Wolbachia*-like sequences are functional. While it is possible that transferred DNA sequences could be expressed if they randomly inserted in a favorable location, it stands to reason that selective pressure would be required to maintain all of the sequences, both coding and regulatory, necessary to

produce a protein, especially one whose expression is spatially and temporally regulated like our putative *Wolbachia* LolC homolog.

The expression of *Wolbachia*-like RNAs and proteins implies, but does not guarantee, function. In order to impact the biology of the recipient, a transferred gene would have to be expressed at the protein level, fold properly, perhaps undergo post-translational modifications, make its way to a useful location (tissue and sub-cellular location) and coordinate with the existing biochemical machinery present in that location. If our putative LolC homolog had been discovered in a genetically tractable organism, *Caenorhabditis elegans* for example, it would be easy to assess its importance based on the impact of gene silencing or deletion. However, determining the exact biochemical role of this protein in the context of the recipient's biochemical landscape could still be a challenge. This is confounded by the fact that we don't have a clear idea of whether other members of the LolCDE transport apparatus are present in *O. flexuosa* or if LolC could have an independent function unrelated to transport.

HGT is a well-characterized mechanism of genome evolution among plant parasitic nematodes. In 1998, the identification of bacteria-like cellulase enzymes in two parasitic cyst nematodes led to the hypothesis that plant parasitism may be facilitated by HGT from soil bacteria [17,18]. Since then, sequencing projects have led to the discovery of potential HGT events from a diverse collection of microbes in many nematode species [19,20,21,22]. Unlike the transfer events we have seen so far in filarial worms, the genes acquired by the plant parasites have obvious roles in the biology of the recipients, allowing them to occupy new ecological niches. More work will be required to determine whether transferred *Wolbachia*-like genes support *Wolbachia*-free filaria in a similar

manner. The next steps in this line of research should include a more thorough analysis of the genomes of *Wolbachia*-free filarial nematodes. This would provide a better inventory of full-length inserts with coding potential and information on their genomic context (i.e. where they were inserted, what genes are nearby, etc.). Better sequence data would, in turn, facilitate more thorough proteomic analyses.

Future sequencing efforts should probably focus on a *Wolbachia*-free species that can be maintained in the laboratory in order to facilitate follow-up studies. Although HGT appears to have been extensive in *O. flexuosa*, there is no way to experimentally determine their significance in this wild-caught species. However, new RNAi protocols may allow us to explore the role of transferred genes in *Wolbachia*-free laboratory models like *A. viteae* as sequences of interest are identified [23,24].

SECTION IV: OTHER CONTRIBUTIONS

Section IV-I: The Characterization of a *Wolbachia*-free species

The characterization of *C. quiscali* (subfamily: Splendidofilariinae) as a *Wolbachia*-free species, described in Appendix 1, should prove useful for several reasons. First, *C. quiscali* is an avian parasite of a previously unexplored subfamily, the Splendidofilariinae [10], which adds to our basic knowledge of the distribution of *Wolbachia* among filarial nematodes. Second, this species is relatively easy to obtain throughout most of the continental United States and Canada due to its prevalence and broad host range. This stands in contrast to other *Wolbachia*-free species that are geographically restricted and/or infect less tractable hosts (e.g., humans, cattle, etc.).

Third, the worms are relatively large in the context of the grackle brain, so they are easy to locate and separate from host tissues (**Fig. 1**).

Aside from providing insight into the evolutionary history of *Wolbachia* and filarial nematodes, assessing the *Wolbachia* status of a wider array of species might provide more opportunities to compare and contrast *Wolbachia*-dependent and independent species. It should also be considered that comparisons including species that are presently *Wolbachia*-free but infected in the distant past (i.e., *O. flexuosa* and *A. viteae*) may result in different findings than studies that include species that were never associated with the endobacteria at all (presumably species outside the Onchocercinae and Dirofilariinae, including the Splendidofilariinae). Therefore, the characterization of *Wolbachia*-free species from distantly related subfamilies would be informative.

Future work should include the examination of species from understudied subfamilies, particularly the Icosiellinae and the Lemdaninae, which have not been examined at all. Genomic surveys along the lines of those described in Chapter 3 of this thesis could be used to determine if any *Wolbachia*-free species outside the Dirofilariinae and Onchocercinae show evidence of HGT that might indicate contact with *Wolbachia* in the distant past. Studies such as these would broaden our understanding of the evolutionary history of *Wolbachia* in filarial nematodes.

Section IV-II: Whole genome amplification for use with filarial nematode DNA

One major factor that impedes the study of filarial nematodes is the inability to obtain sufficient biological material for experimentation or sequencing. Relatively few filarial species can be maintained in the laboratory, and specimens of other species/strains

must be collected from wild animal hosts or infected humans. These specimens are difficult to obtain, and they often consist of a few microfilaria from infected blood rather than adult worms. These samples contain very small amounts of DNA. Our evaluation of a multiple displacement amplification technique for the efficient and faithful replication of filarial DNA will help to overcome this obstacle.

The ability to amplify useful amounts of DNA from a small amount of starting material will facilitate many studies that might not have otherwise been possible. For example, genome amplification could be used to amplify DNA from parasites collected in the field in order to assess regional variability within species and to identify genetic markers associated with tissue tropism, periodicity in peripheral blood, or drug susceptibility.

Section IV-III: Sequencing projects

Relatively little genetic information is available representing the filarial nematodes, especially given the significant medical and economic burden they present. Only one draft genome, that of *Brugia malayi*, has been thoroughly sequenced (estimated 9x coverage), and assembly is still a work in progress [5]. Projects to sequence the genomes of other human pathogens such as *Onchocerca volvulus*, *Loa loa* and *Wuchereria bancrofti* have been initiated, and sequences related to these organisms are presently available in the form of shotgun reads; however, no assemblies or annotations have been reported. Several filarial species are represented in the Genbank EST database, but many of these EST projects (with notable exceptions of *B. malayi* and *O. volvulus*) have been limited in scope (**Table 2**). During the course of this project, we sequenced

mtDNAs, surveyed the genomes of *A. viteae* and *O. flexuosa* and sequenced the adult transcriptome of *O. flexuosa*. These efforts yielded significant additions to the sequence data available from filarial parasites.

Our desire to compare the mitochondria of *Wolbachia*-dependent and independent led to the sequencing and annotation of the mtDNAs of five filarial species. The mtDNA sequences will be useful in future studies of mitochondrial function and other analyses. Mitochondrial genes such as cytochrome oxidase subunit I and the 12s rRNA gene [16,25,26] are commonly used as phylogenetic markers, so these mtDNAs may facilitate better characterization of evolutionary relationships among filarial nematodes. Additionally, the conserved primer sets used to amplify portions of the mtDNAs for cloning and sequencing will make it easy to repeat these studies in other species.

As reported in Chapter 4, 454 sequencing of *O. flexuosa* cDNA produced sequences representing up to 40% of all *O. flexuosa* protein coding genes. We consider this level of coverage to be quite good, particularly given our restriction to one life cycle stage of this wild-caught species. This dataset should prove useful to researchers in our field since it now represents the most thorough examination of expressed genes currently available for a *Wolbachia*-free filarial species. When further transcriptome data becomes available from closely related *Wolbachia*-dependent species (e.g., *O. volvulus* or *Onchocerca ochengi*) our data may facilitate the first global comparison of gene content and expression between sister taxa with discordant *Wolbachia* status.

Section IV-IV: Proteomic analysis of *O. flexuosa*

In the past, proteomic analyses of filarial nematodes have focused on a single laboratory strain of *B. malayi*, an organism with a nearly complete and well-annotated genome that can be used for comparison with mass spectrometry data [5]. Many of these studies were further restricted to the *B. malayi* secretome in order to limit the complexity of the analyzed sample [27,28,29]. Only recently was a thorough analysis performed on total worm lysate from several life cycle stages [30]. The proteomic analysis of *O. flexuosa* presented in Chapter 4 represents the first shotgun proteomic analysis of a *Wolbachia*-free filarial nematode parasite.

The identification of 1,803 proteins from adult worm lysate provided a first look at protein-level gene expression and validated more than 700 peptide translations from the *O. flexuosa* transcriptome. Our analysis also served as a proof-of-principle for the use of a large comparative database containing putative protein sequences from a wide array of species in shotgun proteomic studies of non-model filarial nematodes. The ability to identify proteins based on orthologs in close relatives implies that proteomic studies need not be restricted to model organisms with well-curated genomes, as useful data can be obtained without an extensive species-specific database. In light of this finding, it may be beneficial to turn our attention towards a *Wolbachia*-free species that can be maintained in the laboratory (e.g., *A. viteae*) even if the sequence data available for this species is not as extensive. A laboratory model would be easier to study and would facilitate better follow-up studies on interesting proteins.

SECTION V: CONCLUSION

The major goal of this thesis was to determine why certain species of filarial nematodes are dependent upon *Wolbachia* endobacteria for growth, fertility, and sometimes even survival while other species are not. Although we have yet to determine the exact mechanisms underlying this biologically and medically significant relationship, the data described here provide some interesting clues. While *Wolbachia*-dependent filarial species must lack some capabilities in comparison to their *Wolbachia*-independent counterparts, it is not likely that this deficiency is related to the sequences of their mitochondrial genomes. It is, however, possible that *Wolbachia*-free filarial species were freed from their dependence on the endosymbiont via HGT, as some *Wolbachia*-like sequences present in the nuclear genomes of *Wolbachia*-free filaria are expressed at the RNA and protein levels and are therefore capable of mediating biological functions. It is our hope that the tremendous body of information collected over the course of this project (mtDNA sequences, genomic DNA sequences, transcript sequences, and proteomic data) will facilitate future studies that help to answer the remaining questions about the nature of the fascinating relationship between filarial nematodes and their *Wolbachia* endosymbionts.

REFERENCES

1. Taylor MJ, Bandi C, Hoerauf A (2005) *Wolbachia* bacterial endosymbionts of filarial nematodes. Adv Parasitol 60: 245-284.
2. Werren JH, Baldo L, Clark ME (2008) *Wolbachia*: master manipulators of invertebrate biology. Nat Rev Microbiol 6: 741-751.
3. Bandi C, Anderson TJ, Genchi C, Blaxter ML (1998) Phylogeny of *Wolbachia* in filarial nematodes. Proc Biol Sci 265: 2407-2413.
4. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753-1756.
5. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al. (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317: 1756-1760.
6. Hu M, Gasser RB, Abs El-Osta YG, Chilton NB (2003) Structure and organization of the mitochondrial genome of the canine heartworm, *Dirofilaria immitis*. Parasitology 127: 37-51.
7. Keddie EM, Higazi T, Unnasch TR (1998) The mitochondrial genome of *Onchocerca volvulus*: sequence, structure and phylogenetic analysis. Mol Biochem Parasitol 95: 111-127.
8. Ferri E, Bain O, Barbuto M, Martin C, Lo N, et al. (2011) New Insights into the Evolution of *Wolbachia* Infections in Filarial Nematodes Inferred from a Large Range of Screened Species. PLoS One 6: e20843.
9. Yatawara L, Wickramasinghe S, Rajapakse RP, Agatsuma T (2010) The complete mitochondrial genome of *Setaria digitata* (Nematoda: Filarioidea): Mitochondrial gene content, arrangement and composition compared with other nematodes. Mol Biochem Parasitol 173: 32-38.
10. Anderson RC (2000) Nematode Parasites of Vertebrates: Their Development and Transmission. New York, New York, USA: CABI Publishing.
11. Ballard JW, Melvin RG (2007) Tetracycline treatment influences mitochondrial metabolism and mtDNA density two generations after treatment in *Drosophila*. Insect Mol Biol 16: 799-802.
12. Brand MD, Nicholls DG (2011) Assessing mitochondrial dysfunction in cells. Biochem J 437: 575.
13. Krueger A, Fischer P, Morales-Hojas R (2007) Molecular phylogeny of the filaria genus *Onchocerca* with special emphasis on Afrotropical human and bovine parasites. Acta Trop 101: 1-14.
14. Plenge-Bonig A, Kromer M, Buttner DW (1995) Light and electron microscopy studies on *Onchocerca jakutensis* and *O. flexuosa* of red deer show different host-parasite interactions. Parasitol Res 81: 66-73.
15. Doolittle WF (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14: 307-311.
16. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. Int J Parasitol 34: 191-203.

17. Keen NT, Roberts PA (1998) Plant parasitic nematodes: digesting a page from the microbe book. Proc Natl Acad Sci U S A 95: 4789-4790.
18. Smant G, Stokkermans JP, Yan Y, de Boer JM, Baum TJ, et al. (1998) Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc Natl Acad Sci U S A 95: 4906-4911.
19. Abad P, Gouzy J, Aury JM, Castagnone-Sereno P, Danchin EG, et al. (2008) Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. Nat Biotechnol 26: 909-915.
20. Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, et al. (2008) The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. Nat Genet 40: 1193-1198.
21. Kikuchi T, Cotton JA, Dalzell JJ, Hasegawa K, Kanzaki N, et al. (2011) Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. PLoS Pathog 7: e1002219.
22. Zarowiecki M (2011) Animals learn new tricks from microorganisms. Nat Rev Microbiol 9: 836.
23. Song C, Gallup JM, Day TA, Bartholomay LC, Kimber MJ (2010) Development of an in vivo RNAi protocol to investigate gene function in the filarial nematode, *Brugia malayi*. PLoS Pathog 6: e1001239.
24. Landmann F, Foster JM, Slatko BE, Sullivan W (2012) Efficient in vitro RNA interference and immunofluorescence-based phenotype analysis in a human parasitic nematode, *Brugia malayi*. Parasit Vectors 5: 16.
25. Huang H, Wang T, Yang G, Zhang Z, Wang C, et al. (2009) Molecular characterization and phylogenetic analysis of *Dirofilaria immitis* of China based on COI and 12S rDNA genes. Vet Parasitol 160: 175-179.
26. Yatawara L, Wickramasinghe S, Nagataki M, Rajapakse RP, Agatsuma T (2007) Molecular characterization and phylogenetic analysis of *Setaria digitata* of Sri Lanka based on CO1 and 12S rDNA genes. Vet Parasitol 148: 161-165.
27. Hewitson JP, Harcus YM, Curwen RS, Dowle AA, Atmadja AK, et al. (2008) The secretome of the filarial parasite, *Brugia malayi*: proteomic profile of adult excretory-secretory products. Mol Biochem Parasitol 160: 8-21.
28. Bennuru S, Semnani R, Meng Z, Ribeiro JM, Veenstra TD, et al. (2009) *Brugia malayi* excreted/secreted proteins at the host/parasite interface: stage- and gender-specific proteomic profiling. PLoS Negl Trop Dis 3: e410.
29. Moreno Y, Geary TG (2008) Stage- and gender-specific proteomic analysis of *Brugia malayi* excretory-secretory products. PLoS Negl Trop Dis 2: e326.
30. Bennuru S, Meng Z, Ribeiro JM, Semnani RT, Ghedin E, et al. (2011) Stage-specific proteomic expression patterns of the human filarial parasite *Brugia malayi* and its endosymbiont *Wolbachia*. Proc Natl Acad Sci U S A 108: 9649-9654.

FIGURES

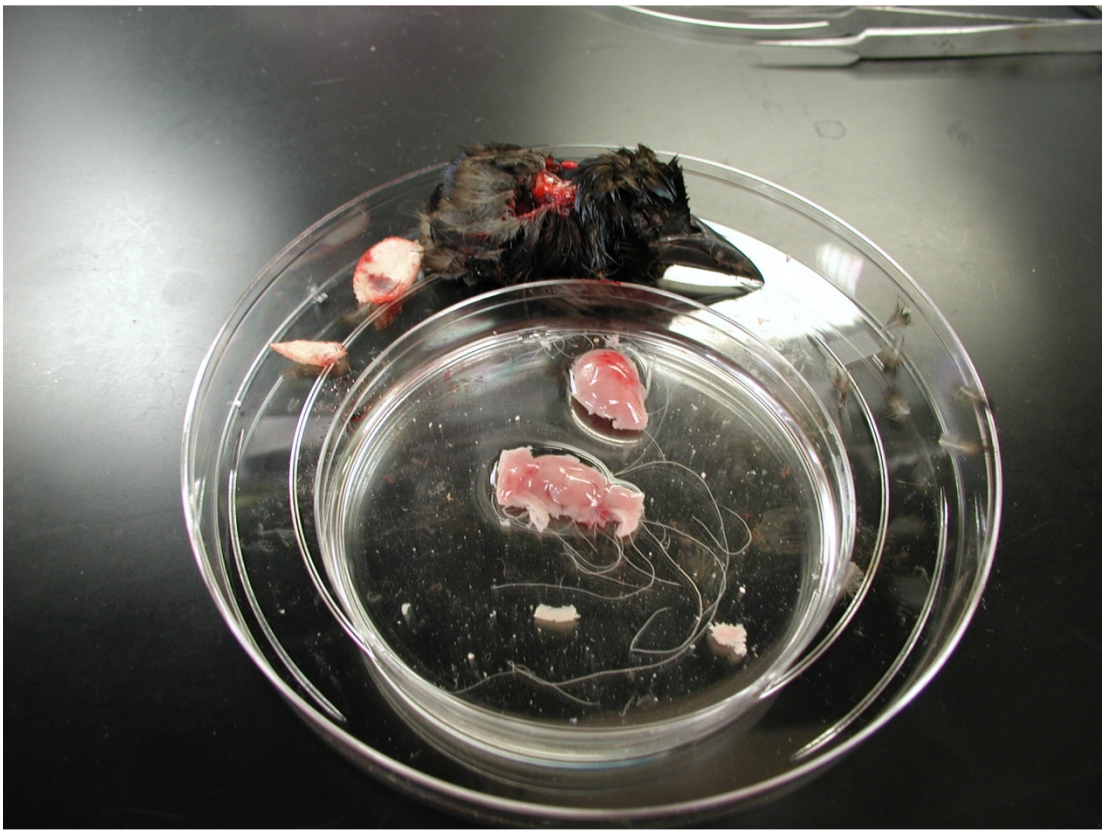


Figure 1: *Chandlerella quiscali* in the common grackle. The worms can be easily identified and dissected from host tissue.

TABLES

Table 1: Project summary

Problem	**Solution**	**Result**	**Future Directions**
Reasons for *Wolbachia*-dependence have not been explained.	Compare *Wolbachia*-dependent and independent species.	No sequence-level differences between the mtDNAs of *Wolbachia*-dependent and independent species.	Compare complete genomes of closely related *Wolbachia*-dependent and independent species pairs.
Reasons for *Wolbachia*-independence have not been explained.	Identify horizontally transferred genes capable of supporting *Wolbachia*-free filarial species.	*A. viteae* and *O. flexuosa* genomes show evidence of HGT from *Wolbachia*. Transferred fragments are expressed at the RNA and protein levels in *O. flexuosa*.	More complete analysis of the genomes and proteomes of *Wolbachia*-free species.
Wolbachia-free filarial worms are difficult to obtain.	Identify *Wolbachia*-free species that are abundant and easy to collect.	*C. quiscali*, a parasite of the common grackle, was shown to lack *Wolbachia*.	Use *C. quiscali* as a model in the study of *Wolbachia*-free filarial nematodes. Assess the *Wolbachia* status of more species.
Scarcity of material representing non-model filarial species.	Evaluate methods for whole genome amplification.	Multiple displacement amplification (MDA) efficiently and accurately amplifies filarial DNA.	Use MDA to obtain sufficient amounts of DNA to characterize scarce material from wild species and strains.
Lack of sequence data representing *Wolbachia*-free species.	Sequence genomes and/or transcriptomes of *Wolbachia*-free species.	Sequenced and annotated the mitochondrial genomes of three *Wolbachia*-free species. Generated 16,814 isogroup and 47,252 high-quality singleton sequences by 454 sequencing of *O. flexuosa* transcriptome.	More thorough analysis of genomes and transcriptomes of *O. flexuosa* and other *Wolbachia*-free species.
Lack of proteomic data representing non-model filaria	Shotgun proteomic analysis of *Onchocerca flexuosa*.	Identified 1,803 proteins from *O. flexuosa* adult worm lysate.	Further proteomic studies of *O. flexuosa* and other filarial species.

Table 2: EST sequences available from various species of filarial nematodes. The number of Sanger ESTs available in GenBank is listed for each species. Information was obtained by searching the NCBI nucleotide EST database (http://www.ncbi.nlm.nih.gov/nucest).

Species	**No. ESTs**
Brugia malayi	26,215
Onchocerca volvulus	14,974
Wuchereria bancrofti	4,874
Loa loa	4,173
Dirofilaria immitis	4,005
Litomosoides sigmodontis	2,699
Onchocerca flexuosa	2,124
Onchocerca ochengi	60
Brugia pahangi	28

APPENDIX 1:

Absence of *Wolbachia* Endobacteria in *Chandlerella quiscali*: an Avian Filarial Parasite

J Parasitol. 2011 Oct 27. doi: 10.1645/GE-2879.1

Samantha N. McNulty[1], Kerstin Fischer[1], Joseph O. Mehus[2], Jefferson A. Vaughan[2], Vasyl V. Tkach[2], Gary J. Weil[1], Peter U. Fischer[1]

[1]Infectious Diseases Division, Department of Internal Medicine, Washington University School of Medicine, Campus Box 8051, 660 S. Euclid Avenue, St. Louis, Missouri 63110.

[2]Department of Biology, University of North Dakota, Grand Forks 58202-9019, USA.

PREFACE

SNM designed and implemented PCR assays to detect *Wolbachia* endobacteria, assisted with *in situ* hybridizations, and wrote the first complete draft of this manuscript. Comments from co-authors and reviewers were incorporated into the published version.

ABSTRACT

Chandlerella quiscali is a filarial nematode parasite of the common grackle (*Quiscalus quiscula*), a widespread bird species found throughout most of North America. Worms collected from wild-caught birds were morphologically identified as *C. quiscali* and tested for the presence of *Wolbachia*, an alphaproteobacterial endosymbiont required for reproduction and maturation by many filarial species. Although various methods, including polymerase chain reaction, in situ hybridization and immunohistology, were used, we were unable to detect evidence of colonization with *Wolbachia*. Due to the widespread distribution of the grackle host, localization within the host, and high prevalence, *C. quiscali* may be among the most easily obtainable of *Wolbachia*-free filarial species. Further studies of *C. quiscali* and other *Wolbachia*-free filarial species may help to clarify the reason(s) that some filarial species require *Wolbachia* but others do not.

INTRODUCTION

Filarial nematodes are a biologically diverse superfamily of parasitic worms that infect a wide array of vertebrates including amphibians, reptiles, birds, mammals, and humans. The causative agents of lymphatic filariasis ([LF], (*Wuchereria bancrofti* and *Brugia malayi*) and onchocerciasis (*Onchocerca volvulus*) infect approximately 150 million people in tropical and subtropical regions [1,2]. New treatments are being sought to prevent and eliminate these infections because they are important causes of long-term disability [1].

Attention has recently shifted to *Wolbachia* endobacteria (Rickettsiales: Anaplasmataceae) as a target for anti-filarial drugs. *Wolbachia* are common reproductive parasites of insects, but they behave as mutualists in many filarial nematode species [3]. Some of the most important filarial pathogens such as *B. malayi*, *W. bancrofti*, *Dirofilaria immitis*, and *O. volvulus*, are dependent on *Wolbachia* endobacteria for growth, fertility, and sometimes even survival [4,5,6,7]. Other filarial species are naturally *Wolbachia*-free and thrive in the absence of a bacterial partner [8,9,10,11,12,13]. The biological mechanisms responsible for this disparity are poorly understood.

Phylogenetic studies may provide insight into the evolutionary history of *Wolbachia*-filaria relationships and improve our understanding of the biological basis of *Wolbachia* dependence in some species. Early surveys focused mainly on medically and economically important, and readily available, parasites of humans and domestic animals. Most of these species are *Wolbachia*-dependent, but most also belong to just 2 of the 10

filarial subfamilies, the Onchocercinae and the Dirofilariinae [8,14]. A recent study by Ferri et al. examined 35 species from 6 filarial subfamilies for *Wolbachia* infection, most of which were *Wolbachia*-free [13]. Their findings suggest that *Wolbachia*-dependent species may be in the minority in many subfamilies and that *Wolbachia* may be entirely absent in filarial parasites of lizards, frogs and birds.

Here, we report results of studies to determine whether *Wolbachia* is present in *Chandlerella quiscali* (Splendidofilariinae), a parasite of the common grackle (*Quiscalus quiscula*) that was first described by Von Linstow in 1904 [15]. Grackles are widespread in North America east of the Rocky Mountains. This makes *C. quiscali* accessible throughout most of the United States and Canada. Polymerase chain reaction (PCR) studies, in situ hybridization and immunohistology were performed to detect evidence of colonization with *Wolbachia* endobacteria, but none was found. Our results suggest *C. quiscali* as a readily available model in the study of *Wolbachia*-free filarial parasites.

MATERIALS AND METHODS

Parasite materials

Adult *C. quiscali* were dissected from the cerebral ventricles of euthanized *Q. quiscula* trapped in North Dakota, USA [16]. Trapping was conducted under USFW Permit MB072162, ND Game & Fish Permit GNF02344201, and University of North Dakota IACUC protocol 0705-3c. Adult *B. malayi* and *Acanthocheilonema viteae* were obtained from experimentally infected Mongolian gerbils (*Meriones unguiculatus*) as previously described [17,18]. Adult *O. volvulus*, adult *Onchocerca flexuosa* and *W. bancrofti* microfilariae (mf) were available from previous studies [19,20,21]. Captive bred field crickets (*Gryllus bimaculatus*) were purchased from a pet shop in Hamburg, Germany.

For PCR studies, harvested worms were frozen in phosphate-buffered saline (PBS) or TRIzol (Invitrogen, Carlsbad, California) at -80 C until use. For in situ hybridization and immunohistology, adult worms and crickets were fixed for 24-72 hr in 4% buffered formaldehyde, embedded in paraffin, and sectioned using standard histological procedures. For morphological examination, adult worms were killed with heated saline, preserved in 70% ethanol, and cleared in lactophenol. Worms were studied using an Olympus BX-51 compound light microscope (Olympus, Tokyo, Japan) equipped with DIC optics and a digital imaging system. Measurements were obtained using Rincon software version 7.1.2 (Imaging Planet, Goleta, California).

DNA isolation

DNA was isolated from *B. malayi, O. volvulus, O. flexuosa, A. viteae* and *C. quiscali* adult worms; *W. bancrofti* mf; and *G. bimaculatus* formalin-fixed, paraffin embedded histological sections. The DNeasy Blood and Tissue Kit (Qiagen, Germantown, Maryland) was used to isolate DNA from adult worms according to the manufacturer's suggested protocol. For *G. bimaculatus,* paraffin was removed from the tissue by 2 30-min incubations with 100% xylene, 2 30-min incubations with 100% ethanol, 2 30-min minute incubations with 75% ethanol, and 2 15-min incubations with 1x PBS. After removing the final PBS wash, tissue was re-suspended in 15 µl 1x PBS and subjected to alkaline lysis. The same alkaline lysis protocol was used to obtain DNA from *W. bancrofti* mf. In brief, tissue was brought to a final volume of 15 µl in sterile water. Fifteen microliters of alkaline lysis buffer (400 mM KOH, 100 mM DTT and 10 mM EDTA) was added, and the mixture was heated to 95 C for 30 min before chilling on ice for 10 min. The solution was neutralized with 15 µl of neutralization buffer (400 mM HCl and 600 mM Tris-HCl pH 7.5). One microliter of lysate or 10 ng purified DNA served as template in PCR reactions.

PCR reactions

Primers used to amplify the 5S ribosomal intergenic spacer were S2 5'-GTTAAGCAACGTTGGGCCTGG-3' and S16 5'-TTGACAGATCGGACGAGATG-3' [22]. Primers used to detect *Wolbachia* endobacteria are reported in **Table I**. *Wolbachia* primers are designed to match portions of the most highly conserved *Wolbachia* genes

[23]. The 5s intergenic spacer was cloned using the TOPO-TA cloning kit for sequencing (Invitrogen) and sequenced.

In situ hybridization

A 424-bp segment of the 16S rDNA gene of the *Wolbachia* endosymbiont of *B. malayi* was amplified from *B. malayi* genomic DNA using the following primers: 16sF 5'-CAGCTCGTGTCGTGAGATGT-3' and 16sR '5-CCCAGTCATGATCCCACTT-3' [24]. Biotinylated antisense and sense (negative control) RNA probes were constructed from this sequence as previously described [24].

For hybridization, 5-μm histological sections were deparaffinized, digested in pepsin HCl for 7 min and hybridized overnight at 60 C in humid chambers with 1 μg of probe in hybridization buffer (50% formamide, 5x SSC, 0.3 mg/ml yeast tRNA, 100 μg/ml heparin, 1x Denhardt's solution, 0.1% CHAPS, and 5 mM EDTA). A stringency wash was performed at 60 C for 30 min, and detection was carried out using the In situ Hybridization Detection System (Dako North America, Carpinteria, California). Sections were incubated with streptavidin-AP conjugate for 20 min at room temperature and with BCIP/NBT for 10-20 min to localize bound probe.

Immunohistochemistry

The alkaline phosphatase anti-alkaline phosphatase technique was used for immunostaining according to the recommendations of the manufacturer (Dako North America) and as described previously [24,25]. A monoclonal antibody directed against *Wolbachia* surface protein (WSP) of *B. malayi* was used as the primary antibody at a

dilution of 1:100. A monoclonal antibody directed against human heat shock protein 60 (Hsp60 LK2, Sigma, St. Louis, Missouri) was used as a positive control at a dilution of 1:5 because our prior studies have shown that it binds to *Wolbachia* and filarial mitochondria.

RESULTS

Based on morphological examination of adult and mf stages (**Fig. 1**), the grackle parasites that were used in this study belong to *Chandlerella*. Details of adult morphology, such as the rounded shape of the anterior end (**Fig. 1A, B**); relative position of the nerve ring (**Fig. 1A**); position of the vulva relative to the end of the esophagus (**Fig. 1A**); presence of characteristic large cells at the junction between the esophagus and the intestine (**Fig. 1E**); and, most importantly, the configuration of the distal portion of the uterus (**Fig. 1A, C**), confirmed that the worms were *C. quiscali*. The distal portion of the uterus does not extend anteriorly beyond the vulva in *C. quiscali*. In contrast, the uterus extends anteriorly well beyond the vulva and then loops posteriorly to reach the vulva in closely related species. Likewise, the length of spicules in males in our material (left spicule 75 μm, right spicule 82 μm) is identical to that of *C. quiscali* and different from the length of spicules in closely related species (**Fig. 1D**) [26]. Microfilariae in the blood were long (182-197 μm) and sheathed, with blunt, very slightly tapered tails (**Fig. 1F, G**), and this corresponds to the description of *C. quiscali* mf reported by Bartlett and Anderson [26]. Localization of adult worms in the grackle brain also supports the identity of these parasites as *C. quiscali*.

DNA samples from worms taken from separate birds were tested by PCR with primer sets targeting *Wolbachia* 16s rRNA and other highly conserved *Wolbachia* genes, including the DNA-directed RNA polymerase (Rpo B/C), translation elongation factor GTPase (FusA), and succinate dehydrogenase (**Fig. 2**). Primer sets previously used to demonstrate the absence of *Wolbachia* in *L. loa* were also tested (data not shown) [11]. All primer sets employed in this study amplified *Wolbachia* sequences from infected

species (crickets, *B. malayi, W. bancrofti*, and *O. volvulus*) but not from *Wolbachia*-free filarial species (*O. flexuosa* and *A. viteae*) [8,12] or *C. quiscali*. The filarial 5s rRNA intergenic spacer was amplified and sequenced as a positive control to ensure the quality of the *C. quiscali* genomic DNA. This sequence was deposited in Genbank under the accession number HM641830.

An in situ hybridization probe designed to bind highly conserved portions of *Wolbachia* 16s rRNA was also used to test for the presence of the endobacteria in *C. quiscali* (**Fig. 3**). This probe is able to detect a wide array of *Wolbachia* strains, including those in *B. malayi* (**Figs. 3A, D, F**) and in the cricket *G. bimaculatus* (**Fig. 3B**). Although dozens of sections were examined, no staining was detected in *C. quiscali* (**Figs. 3C, E**).

Likewise, a monoclonal antibody against a *Wolbachia* surface protein was used to stain the *Wolbachia* in *B. malayi* (**Fig. 4A**) and *G. bimaculatus* (**Fig. 4C**), but this antibody did not produce a signal in *C. quiscali* (**Figs. 4D, E**). A positive control antibody against Hsp60 that labels filarial mitochondria did produce a signal in *C. quiscali* (**Figs. 4B, F**). This positive control indicates that the *C. quiscali* material was intact and suitable for antibody staining.

DISCUSSION

The worms used in this study were identified as *C. quiscali* based on their morphologic features, host species, and localization within the host [16,26,27]. These characteristics place species of the genus *Chandlerella* in the subfamily Splendidofilariinae [27]. *C. quiscali* is 1 of only 2 species of this subfamily to have been formally examined for *Wolbachia* infection [13].

PCR with different primer sets, in situ hybridization, and immunohistology all indicated that *C. quiscali* does not contain *Wolbachia*. The genes targeted by PCR were among the most highly conserved in sequenced *Wolbachia* strains, as indicated by the fact that we were able to amplify *Wolbachia* sequences present in a diverse array of host species including both insects and filarial nematodes. The in situ probe against *Wolbachia* 16s rRNA and the antibody against WSP also detected highly conserved targets in a wide array of filarial and insect *Wolbachia* strains. Similar methods have been used to confirm the *Wolbachia*-free nature of other filarial species, including *Loa loa* and *O. flexuosa* [9,11,12].

Our understanding of the co-evolutionary dynamics of *Wolbachia* in filarial parasites will benefit from wider sampling of parasite species representing different subfamilies of Onchocercidae. To-date, 54 species of Onchocercid species have been examined for *Wolbachia*, of which 26 (48%) harbor *Wolbachia* [13,14,28,29]. Most (72%) of the species examined to-date belong to a single subfamily, Onchocercinae. Because there is good representation within this subfamily, some patterns can be discerned when comparing parasite genera. For example, 13 of 14 (93%) species tested within the genus *Onchocerca* harbor *Wolbachia*, whereas only 1 of 9 (11%) species

tested within the genus *Cercopithifilaria* harbor *Wolbachia*. This difference in proportions is statistically significant difference (Fisher's exact test, p=0.002). The 2 subfamilies that contain most avian filarioid species (Splendidofilarinae and Lemdaninae) have not been extensively examined for *Wolbachia* symbionts. Avian filarioids (approximately 15 genera) may be among the most widely distributed of the filarioid parasites because of their hosts' unique ability to fly, yet only 2 avian species, an unidentified *Aproctella* species [13] and *Chandlerella quiscali*, both Splendidofilarinae, have been studied. Both of these species are *Wolbachia*-free. More species will need to be examined to determine if this is true of all avian filaria.

Future studies should not only include examination of a wider array of filarial species, but a more in-depth examination of species that have already been characterized as *Wolbachia*-free. Our previous studies uncovered evidence of horizontal gene transfer from *Wolbachia* endobacteria to 2 species that are currently *Wolbachia*-free (*O. flexuosa* and *A. viteae*,). This indicates that their ancestors must have been infected with *Wolbachia* [20]. Genomic surveys of other *Wolbachia*-free species like *C. quiscali* will help to determine whether *Wolbachia* was present in their lineage in the past.

Now that *C. quiscali* has been experimentally shown to lack *Wolbachia*-endobacteria, it can be used as a model for studies of *Wolbachia*-free filarial parasites. *C. quiscali* has certain attributes that may make it an attractive model system for such studies. First, its host, the common grackle, is abundant and synanthropic, it is easy to trap, and it does well in captivity. Local prevalence (particularly in forested regions where vector [ornithophilic *Culicoides*] populations are high) can approach 100% [30,31,32]. Moreover, unlike many avian filarids whose adult stages are small and often

difficult to locate within their hosts, *C. quiscali* adults are easily found within the ventricles of the brain. In depth studies of readily available *Wolbachia*-free species like *C. quiscali* may help to determine why some filarial nematodes require *Wolbachia* for fertility and maturation while others do not.

ACKNOWLEDGEMENTS

We thank Dr. Daojun Jiang (Washington University School of Medicine) for providing PCR primers for the amplification of the 16s rRNA in situ probe and Dr. Patrick Lammie (Center for Disease Control and Prevention) for providing the anti-WSP antibody. This work was supported by NIH grants T32-AI007172 to S.N.M and R03-AI092306 to J.A.V. and by Barnes-Jewish Hospital Foundation Grant 01251-060.

REFERENCES

1. (2009) Global programme to eliminate lymphatic filariasis. The Weekly Epidemiological Record 84: 437-444.
2. Remme JHF, Feenstra P, Lever PR, Medici AC, Morel CM, et al. (2006) Tropical Diseases Targeted for Elimination: Chagas Disease, Lymphatic Filariasis, Onchocerciasis, and Leprosy. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M et al., editors. Disease Control Programs in Developing Countries. 2nd ed. New York: Oxford University Press.
3. Werren JH, Baldo L, Clark ME (2008) *Wolbachia*: master manipulators of invertebrate biology. Nature reviews Microbiology 6: 741-751.
4. Bandi C, McCall JW, Genchi C, Corona S, Venco L, et al. (1999) Effects of tetracycline on the filarial worms *Brugia pahangi* and *Dirofilaria immitis* and their bacterial endosymbionts *Wolbachia*. International Journal for Parasitology 29: 357-364.
5. Casiraghi M, McCall JW, Simoncini L, Kramer LH, Sacchi L, et al. (2002) Tetracycline treatment and sex-ratio distortion: a role for *Wolbachia* in the moulting of filarial nematodes? International Journal for Parasitology 32: 1457-1468.
6. Chirgwin SR, Nowling JM, Coleman SU, Klei TR (2003) *Brugia pahangi* and *Wolbachia*: the kinetics of bacteria elimination, worm viability, and host responses following tetracycline treatment. Experimental Parasitology 103: 16-26.
7. Hoerauf A, Nissen-Pähle K, Schmetz C, Henkle-Dührsen K, Blaxter ML, et al. (1999) Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility. The Journal of Clinical Investigation 103: 11-18.
8. Bandi C, Anderson TJC, Genchi C, Blaxter ML (1998) Phylogeny of *Wolbachia* in filarial nematodes. Proceedings of the Royal Society of London Series B, Biological Sciences 265: 2407-2413.
9. Büttner DW, Wanji S, Bazzocchi C, Bain O, Fischer P (2003) Obligatory symbiotic *Wolbachia* endobacteria are absent from *Loa loa*. Filaria Journal 2: 10.
10. Chirgwin SR, Porthouse KH, Nowling JM, Klei TR (2002) The filarial endosymbiont *Wolbachia* sp. is absent from *Setaria equina*. The Journal of Parasitology 88: 1248-1250.
11. McGarry HF, Pfarr K, Egerton G, Hoerauf A, Akue J-P, et al. (2003) Evidence against *Wolbachia* symbiosis in *Loa loa*. Filaria Journal 2: 9.
12. Plenge-Bonig A, Kromer M, Buttner DW (1995) Light and electron microscopy studies on *Onchocerca jakutensis* and *O. flexuosa* of red deer show different host-parasite interactions. Parasitology Research 81: 66-73.
13. Ferri E, Bain O, Barbuto M, Martin C, Lo N, et al. (2011) New Insights into the Evolution of *Wolbachia* Infections in Filarial Nematodes Inferred from a Large Range of Screened Species. PLoS One 6: e20843.
14. Casiraghi M, Bain O, Guerrero R, Martin C, Pocacqua V, et al. (2004) Mapping the presence of *Wolbachia pipientis* on the phylogeny of filarial nematodes: evidence for symbiont loss during evolution. International Journal for Parasitology 34: 191-203.

15. Linstow OFBV (1904) Beobachtungen an Nematoden und Cestoden. Archiv für Naturgeschichte Berlin 70: 297-309.
16. Odetoyinbo JA (1960) Biology of *Splendidofilaria quiscai* (Von Linstow, 1904) N. Comb. (Nematoda: Onchocercindae) [Ph.D. Thesis.]. Ames, Iowa: Iowa State University of Science and Technology.
17. Ash LR, Riley JM (1970) Development of subperiodic *Brugia malayi* in the jird, *Meriones unguiculatus*, with notes on infections in other rodents. The Journal of Parasitology 56: 969-973.
18. Lucius R, Textor G (1995) *Acanthocheilonema viteae*: rational design of the life cycle to increase production of parasite material using less experimental animals. Applied Parasitology 36: 22-33.
19. Fischer PU, Kipp W, Bamuhiga J, Binta-Kahwa J, Kiefer A, et al. (1993) Parasitological and clinical characterization of *Simulium neavei*-transmitted onchocerciasis in western Uganda. Annals of Tropical Medicine and Parasitology 44: 311-321.
20. McNulty SN, Foster JM, Mitreva M, Hotopp JCD, Martin J, et al. (2010) Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. PLoS One 5: e11029.
21. Weil GJ, Kastens W, Susapu M, Laney SJ, Williams SA, et al. (2008) The impact of repeated rounds of mass drug administration with diethylcarbamazine plus albendazole on bancroftian filariasis in Papua New Guinea. PLoS Neglected Tropical Diseases 2: e344.
22. Fischer PU, Buttner DW, Bamuhiiga J, Williams SA (1998) Detection of the filarial parasite *Mansonella streptocerca* in skin biopsies by a nested polymerase chain reaction-based assay. The American Journal of Tropical Medicine and Hygiene 58: 816-820.
23. Holman AG, Davis PJ, Foster JM, Carlow CK, Kumar S (2009) Computational prediction of essential genes in an unculturable endosymbiotic bacterium, *Wolbachia* of *Brugia malayi*. BMC Microbiology 9: 243.
24. Fischer K, Beatty WL, Jiang D, Weil GJ, Fischer PU (2011) Tissue and Stage-Specific Distribution of *Wolbachia* in *Brugia malayi*. PLoS Neglected Tropical Diseases 5: e1174.
25. Rao RU, Huang Y, Fischer K, Fischer PU, Weil GJ (2009) *Brugia malayi*: Effects of nitazoxanide and tizoxanide on adult worms and microfilariae of filarial nematodes. Experimental Parasitology 121: 38-45.
26. Bartlett CM, Anderson RC (1980) Development of *Chandlerella chitwoodae* Anderson, 1961 (Filarioidea: Onchocercidae) in *Culicoides stilobezzioides* Foote and Pratt and *C. travisi* Vargas (Diptera: Ceratopogonidae). Canadian Journal of Zoology 58: 1002-1006.
27. Anderson RC, Bain O (1976) Keys to genera of the order Spirurida. Part 3. Diplotriaenoidea, Aproctoidea and Filarioidea. In: Anderson RC, Chabaud AG, Willmott S, editors. CIH keys to the nematode parasites of vertebrates. Farnham Royal, UK: Commonwealth Agricultural Bureau. pp. 59-116.
28. McLaren DJ, Worms MJ, Laurence BR, Simpson MG (1975) Micro-organisms in filarial larvae (Nematoda). Transactions of the Royal Society of Tropical Medicine and Hygiene 69: 509-514.

29. Taylor MJ, Hoerauf A (1999) *Wolbachia* bacteria of filarial nematodes. Parasitology Today 15: 437-442.
30. Granath WO (1980) Fat of the wild avian filarial nematode *Chandlerella quiscali* (Onchocercidae: Filarioidea) in the domestic chicken. Poultry Science 59: 996-1000.
31. Johnson AA (1984) Helminths of Common Grackles (*Quiscalus quiscula-versicolor*, Vieillot) in Central Arkansas. Arkansas Acedemy of Science Proceedings 38: 53-55.
32. Welker GW (1962) Helminth Parasites of the Common Grackle, *Quiscalus quiscula versicolor* Veillot in Indiana: Ohio State University.

FIGURES

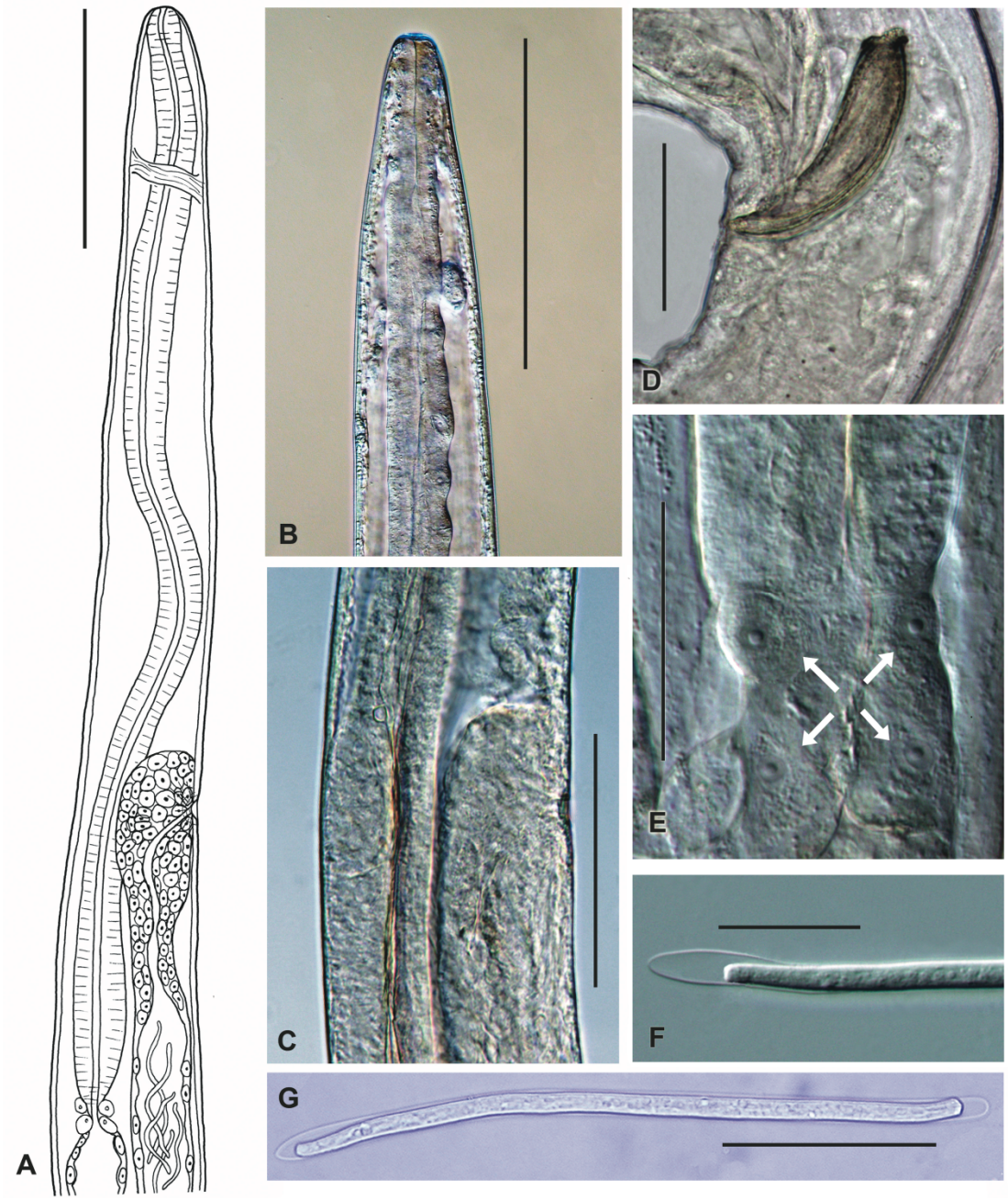


Figure 1: Morphological characteristics of *C. quiscali*. A line drawing of the anterior end of the female demonstrates the relative position of the vulva, esophagus and nerve ring (**A**); the anterior end of *C. quiscali* (**B**); the region of the vulva (**C**; note that the uterus does not make a loop anterior to the vulva); the posterior end of male with spicules (**D**); the of junction between esophagus and intestine (**E**) with characteristic large cells indicated with arrows; and sheathed microfilaria of *C. quiscali* from blood of common grackle (**F**,**G**). Bars = 250 μm (**A, B**), 150 μm (**C**), 50 μm (**D, E, G**), and 25 μm (**F**).

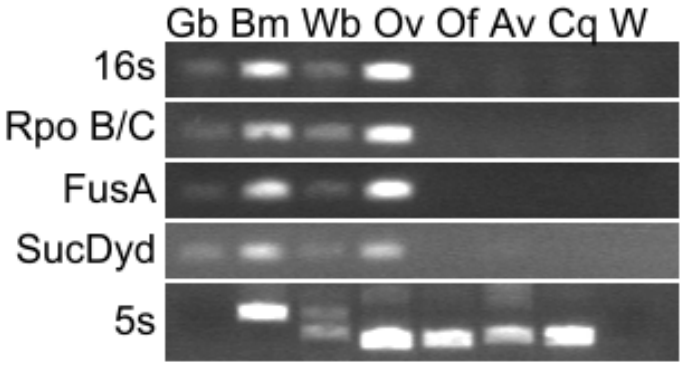


Figure 2: Results of a PCR-based search for evidence of colonization with *Wolbachia* endobacteria in various species. Template species are designated as follows: Gb, field cricket *Gryllus bimaculatus*; Bm, *B. malayi*; Wb, *W. bancrofti*; Ov, *O. volvulus*; Of, *O. flexuosa*; Av, *A. viteae*; Cq, *C. quiscali*; W, water only (a no template control).

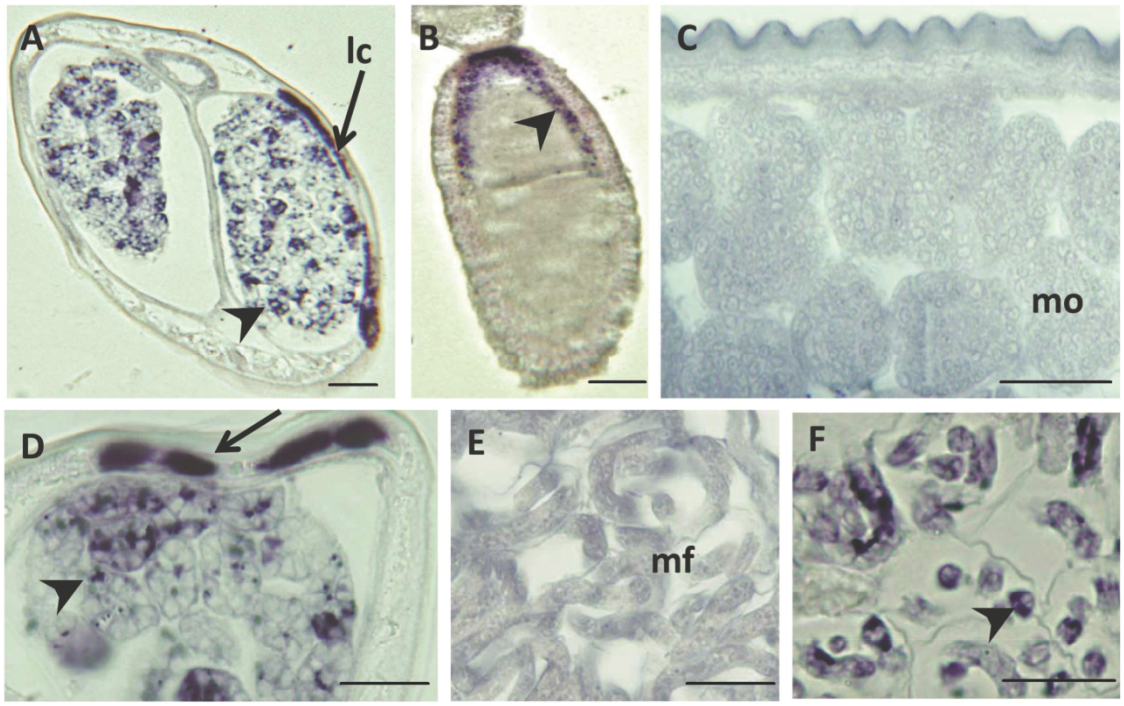


Figure 3: In situ hybridization using a probe for the 16s rRNA of *Wolbachia* endobacteria. The 16s probe bound to *Wolbachia* in the lateral chords (**A, D**), early (**A**) and late (**D**) morulae, and microfilariae (**F**) of *B. malayi*. It also bound *Wolbachia* in the developing eggs of the cricket *G. bimaculatus* (**B**). No staining was detected in the late morulae (**C**) or microfilaria (**E**) of *C. quiscali*. Bars = 25 μm. Species are designated within the figure as follows: Bm, *B. malayi*; Gb, *G. bimaculatus*; Cq, *C. quiscali*.

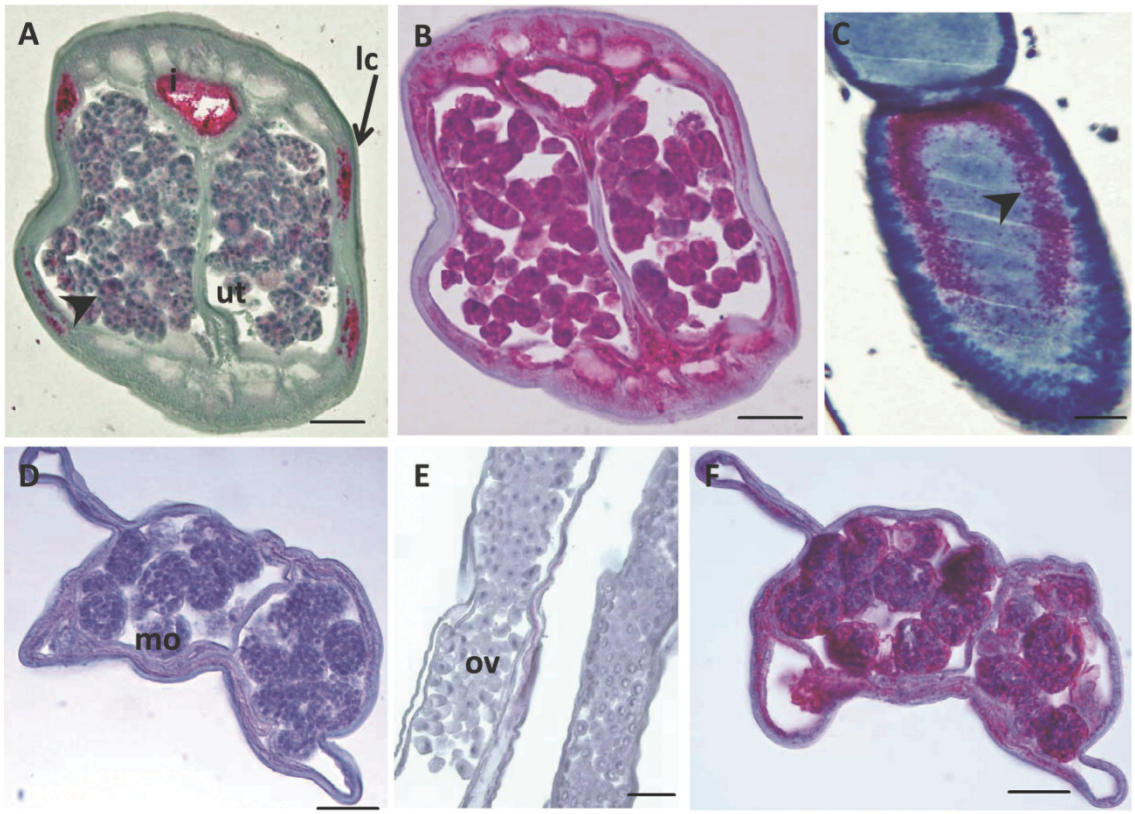


Figure 4: Immunohistological detection of *Wolbachia* surface protein (WSP). WSP antibody labeled *Wolbachia* in the lateral chords (lc) and morula stage embryos in the uterus (ut) of *B. malayi* (**A**) and in developing eggs of *G. bimaculatus* (**C**). No staining was detected in morula stage (mo) embryos (**D**) or the ovaries (ov) (**E**) of *C. quiscali*. A positive control Hsp60 antibody labeled mitochondria in both *B. malayi* (**B**) and *C. quiscali* (**F**). Bars = 25 µm. Species are designated within the figure as follows: Bm, *B. malayi*; Gb, *G. bimaculatus*; Cq, *C. quiscali*.

TABLES

Table I: PCR primer sequences used to detect *Wolbachia* endobacteria (5'-3')

16s Universal	CCAGTGGCGAAGGCGTCTAT	CCCCGTCAATTCCTTTGAGTTT
RpoB/C	TTCTGGCTCTGGTGCTGTAG	AACCTTGCCAAGACATAAAAGC
FusA	GATGGTGCAGCTTCTATGGAT	GCAACTCCATCAAATACAGCAA
SucDyd	CAGGTGGATATGGACGTGTT	CCATATATTCCTGTTGGATGAAA

*SucDyd, succinate dehydrogenase

APPENDIX 2:

Brugia malayi: Whole Genome Amplification for Genomic Characterization of Filarial Parasites

Exp Parasitol. 2008 Jun;119(2):256-63. doi: 10.1016/j.exppara.2008.02.003

Samantha N. McNulty[a], Gary J. Weil[a], Michael Heinz[b], Seth D. Crosby[b] and Peter U. Fischer[a]*

[a] Infectious Diseases Division, Washington University School of Medicine, Campus Box 8051, 660 S. Euclid Avenue, St. Louis, MO 63110, USA

[b] Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA

PREFACE

SNM performed all experiments except the microarray hybridization and wrote the first complete draft of this manuscript. Comments from co-authors and reviewers were incorporated into the published version.

ABSTRACT

Genetic characterization of field isolates and clinical specimens of filarial nematodes is often limited by a shortage of DNA; therefore, we evaluated a multiple displacement amplification (MDA) based whole genome amplification method. The quality of amplified DNA was examined by conventional PCR, real-time PCR, and DNA hybridization. MDA of 5.0 ng of adult *Brugia malayi* DNA and one-fifteenth of the DNA isolated from a single microfilaria resulted in 6.3 μg and 4.2 μg of amplified DNA, respectively. Amplified DNA was equivalent to native genomic DNA for hybridization to *B. malayi* BAC library clones or to an oligonucleotide microarray with approximately 18,000 filarial DNA sequences. MDA is useful for whole genome amplification of filarial DNA from very small amounts of starting material. This technology will permit detailed studies of genetic diversity that were not previously feasible.

INTRODUCTION

Filarial nematode parasites cause serious illnesses such as elephantiasis and "river blindness". With some 150 million people infected and more than 1 billion at risk of infection, these disabling diseases affect the lives of millions and retard economic growth in the developing world.

The drugs that are currently used in control programs are most active against larval (transmissible) stages of the parasite; the drugs kill microfilariae and temporarily sterilize adult worms [1]. Mass drug administration in some endemic areas has reduced the infection and disease rates associated with onchocerciasis and lymphatic filariasis (LF); however, these programs would be more effective if they could employ drugs that effectively kill adult worms. Better tools are needed for identifying new drug targets, for understanding mechanisms of action of existing antifilarial drugs, and for assessing the emergence of drug resistance. Research efforts in this area are severely impeded by a lack of parasite material, as only zoophilic strains of *B. malayi* can be maintained in the laboratory. Studies of other strains or species require isolation of parasite material from human blood, tissues, or insect vectors. Tiny amounts of DNA isolated from these samples are inadequate for thorough genetic analyses. Whole genome amplification (WGA) procedures offer a potential means of overcoming this obstacle.

Various WGA protocols have been described, and several commercial kits for WGA are available. Among these techniques, multiple displacement amplification (MDA) is a particularly attractive option [2,3]. Based on bacterial rolling circle amplification, this method uses a high fidelity Φ29 polymerase and random hexamer primers in an isothermal reaction to non-specifically amplify entire genomes. MDA has

been able to successfully amplify the genomes of many organisms, including mammals, protozoa and bacteria, and several groups have reported that it can produce useful quantities of DNA from template isolated from a single bacterial cell [4]. Of particular interest, MDA has also been used to amplify the genome of *Wolbachia* bacteria in *Drosophila* [5] or to improve the detection of *Wolbachia* in mites [6]. Insect *Wolbachia* are closely related to *Wolbachia* bacteria in filarial nematodes.

The primers included in commercially available MDA reaction kits were designed and optimized for use with human DNA, which has a balanced GC content. In contrast, filarial nematode genomes are highly AT rich (>70%) [7,8], so it is not clear whether this method would provide adequate amplification and unbiased coverage of filarial DNA using primers included with commercial kits. MDA was recently reported to work well with the AT rich *Plasmodium falciparum* genome; however, the filarial genomes are approximately 40 times larger than those of *Plasmodia* [2,8,9,10]. Because quality is as important as quantity, a through analysis of MDA-based amplification of filarial DNA was required before this technique could be used for molecular characterization of filarial parasite material from clinical isolates.

In this study, we amplified DNA from *B. malayi* by MDA and rigorously assessed the quality of amplified DNA (aDNA) by a number of different methods to demonstrate that MDA can be used for representative, high fidelity amplification of filarial nematode genomes.

MATERIALS AND METHODS

Parasite material

Adult *B. malayi* (TRS strain) were obtained from infected jirds (*Meriones unguiculatus*) provided by the NIAID / NIH Filariasis Research Reagent Repository Center (FR3, http://www.filariasiscenter.org/). Microfilariae of the same *B. malayi* strain were isolated from infected jird blood by filtration. Parasites were stored at -20ºC prior to use. Filters containing a gridded genomic BAC library of *B. malayi* were obtained from FR3 [11].

DNA isolation

Genomic DNA was isolated from adult worms using the DNeasy Tissue Kit (Qiagen, Valencia CA), and eluted from the column in 400 µL sterile water. In order to isolate DNA from single MF, MF in suspension were diluted in water to a concentration of approximately 10 MF/ 20 µL. Single MF were then removed from the suspension with a pipette under 40X magnification and added to 5 µL 1x PBS and 5 µL of alkaline lysis buffer (400 mM KOH, 100 mM DTT & 10 mM EDTA). Tubes were heated for 25 min at 95ºC and chilled on ice for 10 minutes before adding 5 µL of neutralization buffer (400 mM HCl, 600 mM Tris-HCl, pH 6.0). Lysis was verified by q-PCR targeting the *B. malayi* 18S rDNA sequence (**Table 1**).

Genome amplification

Genomic DNA was amplified according to the protocol given in the GenomiPhi DNA Amplification Kit (GE Healthcare, Buckinghamshire, UK). Briefly, *B. malayi* template

DNA was diluted to the desired concentration and 1ul was mixed with 9 μL of sample buffer, heated to 95ºC for 3 min and chilled on ice for 5 min. For each reaction, 1 μL of enzyme mix was added to 9 μL of reaction buffer on ice. After cooling, the reaction mixture was combined with the cold template mixture and incubated at 30ºC for at least 18 hours. Reactions were stopped by incubation at 65ºC for 10 min and amplified DNA was purified by ethanol precipitation.

Quantitative Real-Time PCR

Two targets, an abundant tandem repeat (*Hha*I repeat) and the *B. malayi* 18S ribosomal RNA gene (18S rDNA) were used for quantitative real-time PCR (q-PCR). The *Hha*I reaction targeted the entire 320 bp repeat and was carried out as previously described [12]. The presence of the *B. malayi* 18S rDNA sequence was assessed in a Taqman assay that amplifies a 120 bp fragment (**Table 1**). SybrGreen (Applied Biosystems, Foster City, CA) q-PCR was used to compare the relative abundance of single copy and low copy number genes [10] in aDNA and gDNA. Gene targets and primer sequences are reported in **Table 2**. Single copy genes are those found in only one scaffold of the assembled portion of the *B. malayi* genome (71 Mb of estimated 90-95 Mb genome). Low copy number genes are those that fall within the 17.5 Mb of unassembled orphan contigs.

Conventional PCR

Conventional PCR was performed using Platinum Taq HIFI polymerase (Invitrogen, Carlsbad CA, USA) according to the protocol recommended by the manufacturer. The primers sets used were designed from a *B. malayi* scaffold (TIGR assembly 13313, clone

name CA:1383333) which contains established markers from the X chromosome [13]. Different sets of primers were selected to produce amplicons ranging from 1.5 to 3.5 kb (**Table 1**).

BAC Library Hybridization

BAC library hybridization was performed as described by Foster and co-workers [11]. Unsheared genomic (500 ng) and amplified DNA (1000 ng) was biotin labeled using the NEBlot Phototope Kit (New England Biolabs, Ipswich, MA) according to the random primer labeling protocol. The biotinylated probe was ethanol precipitated and tested for labeling efficiency. Identical 6.5 x 11 cm sections of the BAC library filters were hybridized overnight with the remainder of each probe. These membranes were developed using the standard protocol and reagents included in the Phototope Star Detection Kit (NEB).

Microarray DNA Hybridization

Microarray studies used a second-generation filarial microarray developed by the filarial microarray consortium. This array contains 65mer oligonucleotides derived from 15,412 *B. malayi* clusters, 1,016 *Onchocerca volvulus* clusters, 872 *Wuchereria bancrofti* clusters and 803 genes of *Wolbachia* endobacteria from *B. malayi*. The filarial microarray was obtained from the Washington University microarray core facility via FR3 (http://www.filariasiscenter.org).

Genomic *B. malayi* DNA was diluted to a final concentration of 500 ng/μL and used as a reference control sample. For aDNA, the products of six MDA reactions were pooled, ethanol precipitated, and re-suspended in sterile water to achieve a final

concentration of 500-1,000 ng/μL. aDNA and gDNA were both sonicated for 8 min to obtain fragment sizes of 200 to 1,000 bp as verified by agarose gel electrophoresis. Fragmented DNA was labeled with Cy3 or Cy5 using the Micromax ASAP labeling kit (Perkin Elmer, Waltham, MA) with two modifications to the manufacturer's instructions: (1) The labeling mix consisted of 15µL of DNA plus water, 4µL of ASAP labeling buffer, and 1µL of ASAP cy-dye reagent. (2) DNA was purified after labeling using DNA Clean & Concentrator - 5 columns (Zymo Research, Orange, CA) according to the manufacturer's protocol. Five microgram of each labeled DNA was suspended in 26µL water and added to 26µL of formamide-based hybridization buffer (vial 7-Genisphere Inc., Hatfield, PA). Hybridizations were performed as previously described [14]. Briefly, hybridization solution was heated to 95ºC for 5 minutes and then allowed to cool to room temperature. 48µL of hybridization solution was then added to the microarray under a supported glass cover slip (Erie Scientific, Portsmouth, NH) at 43ºC for 16-20 hours at high humidity. Slides were gently submerged into 2X SSC, 0.2% SDS for 11 minutes, transferred to 2X SSC for 11 min., transferred to 0.2X SSC for 11 minutes, and then spun dry by centrifugation. All washes were performed at room temperature. The arrays were treated with Dyesaver (Genisphere) to prevent fluorophore degradation. Slides were scanned on a ScanArray Express HT scanner (Perkin Elmer) to detect Cy3 and Cy5 fluorescence. Laser power was kept constant for Cy3/Cy5 scans and PMT was varied for each experiment based on optimal signal intensity with lowest possible background fluorescence. A low setting PMT scan was also performed to recover signal from saturated elements. Gridding and analysis of images was performed using ScanArray v3.0 (Perkin Elmer). Three biological replicates (aDNA or gDNA) and at least 2

technical replicates (including dye flip) were used in this assay. The microarray experiment followed the MIAME protocol (http://www.mged.org/Workgroups/MIAME/miame.html).

RESULTS

Yield of amplified DNA

Multiple displacement amplification (MDA) was evaluated for its ability to effectively and faithfully amplify the genome of *B. malayi*. We used 5.0 ng of genomic DNA (gDNA) from adult worms as template for this part of the study, as this is roughly equivalent to the amount of DNA that can be obtained from five microfilariae. This is a number that which can be easily obtained from clinical samples. MDA reactions (n = 5) produced an average of 6.3 ± 1.4 μg of DNA (1260-fold amplification). An aliquot (about 6%) of the DNA isolated from a single *B. malayi* MF was also used as template for MDA, and this yielded approximately 4.2 μg per reaction. No-template control reactions commonly produce up to 6 μg aDNA because the amplification procedure is not sequence-specific and working conditions are never completely free of environmental DNA (e.g. bacterial DNA in the enzyme mixture or DNA from the researcher performing the reaction). PCR analysis of negative control products confirmed the absence of filarial DNA sequences in all experiments.

Quality of amplified DNA assessed by real-time PCR

Since MDA can amplify any DNA present in the reaction mixture (including contaminants), q-PCR was used to ensure that amplified DNA is representative of *B. malayi*. Primers and probes targeting two repeated DNA sequences (18s rDNA and the *HhaI* repeat sequences) were used to estimate the equivalence of aDNA to unamplified gDNA template. Cycle threshold (C_T) values obtained with aDNA were consistently slightly higher than those obtained with gDNA. When we used 10 ng of gDNA as

template, the C_T values for aDNA are on average 1.4 cycles (range 0.79 – 1.72) higher for both targets (**Fig. 1**). Since the difference in C_T value corresponds to a $2^{(aDNA\ Ct - gDNA\ Ct)}$ difference in target abundance, the amount of filarial sequence in aDNA should be around 2.6 fold ($2^{1.4}$) less than those present in the same amount of gDNA; however, this trend was not statistically significant (Mann-Whitney U test, $P > 0.05$). If we assume that the average yield of an MDA reaction is roughly 5.0 μg, we calculate that the target sequences in the original 10 ng of gDNA template were amplified approximately 192-fold. Because these q-PCRs targeted multi-copy loci, we followed this study with six more q-PCR assays targeting single-copy or low copy number loci. Amplification patterns were remarkably similar for the six genes tested over the entire dilution range. q-PCR results from these reactions indicate an average C_T difference of 1.46 ± 1.40 between aDNA and gDNA sequences (**Fig. 2**).

Further experiments tested the performance of MDA reaction with lower template concentrations alone and mixed with DNA from different species. q-PCR results showed that *B. malayi* sequences were detected in as little as 1 pg of aDNA made with 0.5 ng of gDNA template and in 1 pg of aDNA made with 10 ng of a 1:1 mixture of human and *B. malayi* genomic template (**Fig. 3**). Target sequences were also detected by q-PCR in as little as 10 pg of total aDNA amplified with template isolated from a single microfilaria. With a single amplification step, approximately 60-90 µg aDNA can be obtained from one MF. q-PCR results indicate that the actual *B. malayi* sequence content in this product is equivalent to between 0.6 and 0.9 µg of *B. malayi* gDNA.

Size of amplified DNA assessed by electrophoresis and conventional PCR

In order to characterize aDNA produced by MDA, we used agarose gel electrophoresis to compare the size of fragments found in aDNA to those of gDNA. Although aDNA contained a wide range of fragment sizes, the largest aDNA fragments were comparable in length to the fragments found in gDNA extracted using a commercial column method (**Fig. 4a**). The smearing seen in the aDNA is a result of the random priming method used in MDA reactions. Random primers do not allow for the amplification of the entire length of each gDNA fragment, so a greater variety of fragment sizes is expected. Presumably, this should not affect the coverage of the MDA reaction, as the template gDNA contains many copies of the genome represented in many different fragments.

To assess the ability of aDNA to substitute for gDNA in conventional PCR reactions, several primer sets were used to amplify fragments of various sizes from both gDNA and aDNA template. **Figure 4b** shows PCR products obtained with primers directed to different sequences in aDNA scaffold that is present in the *B. malayi* X chromosome. Products ranged in size from 1.5 – 3.5 kb, demonstrating that even fairly large fragments can be amplified from as little as 1 ng of aDNA template.

Coverage of the *B. malayi* genome

To determine the coverage of the aDNA product, we hybridized both gDNA and aDNA to a high density membrane array of an 18,000-clone *B. malayi* BAC library that represents the entire *B. malayi* genome (both coding and non-coding sequences) [11]. Only 500-1000 ng of DNA was used to probe the blots in order to avoid saturation. We

adjusted hybridization conditions to account for the difference in *B. malayi* sequence content between the gDNA and aDNA samples used in this experiment by adding roughly twice the amount of aDNA probe to the filter as gDNA probe. Despite this adjustment, q-PCR results showed that the gDNA probe contained 1.7 times more *B. malayi* 18s rDNA than the aDNA probe (data not shown). However, the hybridization patterns are remarkably similar (**Fig. 5**). Comparisons of patterns obtained with the two probes show greater than 90% agreement between the two blots. Variable signal intensities on the blot are probably caused by variations in the copy number of different genomic DNA sequences. These hybridization results suggest that MDA provided adequate and unbiased amplification of the *B. malayi* genome.

Microarray analysis of amplified B. malayi DNA

As a more stringent test of the coverage of the aDNA product, we compared *B. malayi* aDNA to that of gDNA by hybridizing both samples to a filarial microarray containing 18,103 oligonucleotide probes representing protein coding sequences from *B. malayi*, *O. volvulus, W. bancrofti* and *Wolbachia* endobacteria of *B. malayi*. On average, gDNA target sequences showed detectable hybridization signals (signal minus background > 200) with 95% of the *B. malayi* probes while aDNA sequences showed hybridization to 93% of these probes (**Table 3**). As expected, both gDNA and aDNA showed significantly lower hybridization rates for *O. volvulus* and *W. bancrofti* probes. However, both gDNA and aDNA hybridized well with *B.m. Wolbachia* sequences. This shows that MDA efficiently amplified *Wolbachia* DNA along with the parasite chromosomal DNA. Slightly fewer *Wolbachia* flags were labeled with the aDNA, but this

is consistent with our other results as aDNA contains slightly less target sequence. However, no statistically significant differences were observed between the percent hybridization of gDNA and aDNA regardless of the species from which the probes were derived ($P > 0.05$). No *B. malayi* DNA was detected by q-PCR in aDNA produced in the absence of *B. malayi* template, but amplification products from "no template" control reactions hybridized with 5 to 16% of the 65mers on the microarrays. Since working conditions were not DNA-free, it is likely that amplified human, bacterial or other nonspecific DNA contained sequences that were similar enough to filarial sequences to hybridize to some of the oligonucleotides on the microarray. Increasing the stringency of hybridization might reduce this background. Taken together, the hybridization results indicate that the coverage of *B. malayi* aDNA was very similar to that of unamplified gDNA.

DISCUSSION

The experimental results reported here show that MDA is a useful method for faithfully amplifying a complex nematode parasite genome in an unbiased manner. We were able to generate several micrograms of aDNA product with a very small amount of *B. malayi g*DNA template (5.0 ng). Detailed characterization studies showed that aDNA had high quality marks and excellent coverage.

We chose to use a small amount of template for our yield experiments in order to approximate the amount of filarial DNA that can be easily obtained from field samples. MDA reactions typically produced approximately 4-6 μg of product regardless of the amount of template that was added to the reaction. This should not be surprising, because any DNA present in the reaction mixture (e.g. DNA contaminants in the enzyme mixture or other environmental sources) will be amplified until the primers and dNTP's are exhausted. Therefore, the reaction will proceed to completion at approximately 5μg of product regardless of template abundance. It is not the absolute amount of product, but the proportion of aDNA that matches the template that depends to some extent on the amount of template used [7,15]. This accounts for the high yield of the NTC reactions.

Since Filarial DNA sequences are very AT rich, it was necessary to test various template mixtures and concentrations to demonstrate that there was no bias in the amplification reaction. The finding that MDA effectively amplified small amounts of *B. malayi* DNA in the presence of human DNA is very important, as parasite material isolated from clinical samples is often contaminated with host sequences. It was also striking that the MDA reaction was able to produce approximately 0.9 μg of representative filarial DNA with template isolated from a single MF. Thus the MDA

reaction should be able amplify enough genetic material from field samples to identify filarial species, strains, and single nucleotide polymorphisms in individual parasites or in populations of parasites in infected humans. It should also be possible to amplify enough DNA from parasites in infected mosquitoes to type endemic strains without having to take blood from humans.

For genetic characterization of parasites, aDNA needs to be representative of all sequences found in the un-amplified template. These studies showed that aDNA was comparable to gDNA in *B. malayi* sequence coverage by both BAC clone and microarray hybridization methods. Incidentally, this is the first report of successful hybridization of gDNA to a filarial microarray. Limited availability of parasite DNA forced past studies of diversity in filarial parasites to focus on few loci. In contrast, microarray hybridization can provide information on many thousands of genetic markers in a single experiment. The combined use of WGA and DNA hybridization to oligonucleotide microarrays will facilitate detailed genetic studies of field isolates of filarial parasites that were not feasible before.

MDA based genome amplification may also become a useful tool for monitoring parasite populations for development of drug resistance. Resistance to albendazole and ivermectin (widely used in the Global Programme to Eliminate Lymphatic Filariasis for mass drug administration) is already common in veterinary parasites [16,17]. No clear evidence for clinical drug resistance has been shown for lymphatic dwelling filarial parasites to date. However, mutations associated with benzimidazole resistance have been detected in *W. bancrofti*, and responses to treatment are variable [18,19]. The paucity of available anthelmintic drugs makes the threat of drug resistance particularly worrisome,

and MDA-based whole genome amplification in combination with microarray analysis could be very useful tools for tracking and characterizing drug resistance in filarial worms.

Yet another interesting aspect of our microarray data is the fact that both *B. malayi* gDNA and aDNA bound better to probes representing *W. bancrofti* than probes representing *O. volvulus*. This supports the results of previous phylogenetic studies that have indicated that *Brugia* species are more closely related to *Wuchereria* than to *Onchocerca* [20]. In light of this, DNA hybridization to microarrays holds great promise as a tool for phylogenetic studies of filarial parasites. Microarray based studies of aDNA would require much less genomic DNA material than traditional comparative genome hybridization methods while still allowing thorough analyses.

In summary, this project has demonstrated the success of MDA-based whole genome amplification in replicating the AT rich genome of a filarial nematode, *B. malayi*. This technique is a valuable addition to the toolbox of molecular techniques that can be employed in the study of helminth parasites.

ACKNOWLEDGEMENTS

The authors would like to thank the Filariasis Research Reagent Repository (FR3) and the filarial microarray consortium for providing parasite materials and microarrays used in this study.

REFERENCES

1. Ottesen EA (2006) Lymphatic filariasis: Treatment, control and elimination. Adv Parasitol 61: 395-441.
2. Carret CK, Horrocks P, Konfortov B, Winzeler E, Qureshi M, et al. (2005) Microarray-based comparative genomic analyses of the human malaria parasite *Plasmodium falciparum* using Affymetrix arrays. Mol Biochem Parasitol 144: 177-186.
3. Bashiardes S, Salame N, Patsalis PC (2006) Evaluation of whole-genome amplification using multiple-displacement amplification of a limited number of cells. Clin Chem Lab Med 44: 1158-1160.
4. Raghunathan A, Ferguson HR, Jr., Bornarth CJ, Song W, Driscoll M, et al. (2005) Genomic DNA amplification from a single bacterium. Appl Environ Microbiol 71: 3342-3347.
5. Mavingui P, Van VT, Labeyrie E, Rances E, Vavre F, et al. (2005) Efficient procedure for purification of obligate intracellular *Wolbachia pipientis* and representative amplification of its genome by multiple-displacement amplification. Appl Environ Microbiol 71: 6910-6917.
6. Jeyaprakash A, Hoy MA (2004) Multiple displacement amplification in combination with high-fidelity PCR improves detection of bacteria from single females or eggs of *Metaseiulus occidentalis* (Nesbitt) (Acari: Phytoseiidae). J Invertebr Pathol 86: 111-116.
7. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A 99: 5261-5266.
8. Ghedin E, Wang S, Foster JM, Slatko BE (2004) First sequenced genome of a parasitic nematode. Trends Parasitol 20: 151-153.
9. Gardner MJ, Hall N, Fung E, White O, Berriman M, et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419: 498-511.
10. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al. (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. Science 317: 1756-1760.
11. Foster JM, Kamal IH, Daub J, Swan MC, Ingram JR, et al. (2001) Hybridization to high-density filter arrays of a *Brugia malayi* BAC library with biotinylated oligonucleotides and PCR products. Biotechniques 30: 1216-1218, 1220, 1222 passim.
12. Rao RU, Weil GJ, Fischer K, Supali T, Fischer P (2006) Detection of *Brugia* parasite DNA in human blood by real-time PCR. J Clin Microbiol 44: 3887-3893.
13. Underwood AP, Supali T, Wu Y, Bianco AE (2000) Two microsatellite loci from *Brugia malayi* show polymorphisms among isolates from Indonesia and Malaysia. Mol Biochem Parasitol 106: 299-302.
14. Li BW, Rush AC, Crosby SD, Warren WC, Williams SA, et al. (2005) Profiling of gender-regulated gene transcripts in the filarial nematode *Brugia malayi* by cDNA oligonucleotide array analysis. Mol Biochem Parasitol 143: 49-57.
15. Bergen AW, Qi Y, Haque KA, Welch RA, Chanock SJ (2005) Effects of DNA mass on multiple displacement whole genome amplification and genotyping performance. BMC Biotechnol 5: 24.

16. Wolstenholme AJ, Fairweather I, Prichard R, von Samson-Himmelstjerna G, Sangster NC (2004) Drug resistance in veterinary helminths. Trends Parasitol 20: 469-476.
17. Prichard RK (1990) Anthelmintic resistance in nematodes: extent, recent understanding and future directions for control and research. Int J Parasitol 20: 515-523.
18. Schwab AE, Boakye DA, Kyelem D, Prichard RK (2005) Detection of benzimidazole resistance-associated mutations in the filarial nematode *Wuchereria bancrofti* and evidence for selection by albendazole and ivermectin combination treatment. Am J Trop Med Hyg 73: 234-238.
19. Eberhard ML, Lammie PJ, Dickinson CM, Roberts JM (1991) Evidence of nonsusceptibility to diethylcarbamazine in *Wuchereria bancrofti*. J Infect Dis 163: 1157-1160.
20. Casiraghi M, Bordenstein SR, Baldo L, Lo N, Beninati T, et al. (2005) Phylogeny of *Wolbachia pipientis* based on gltA, groEL and ftsZ gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. Microbiology 151: 4015-4022.

FIGURES

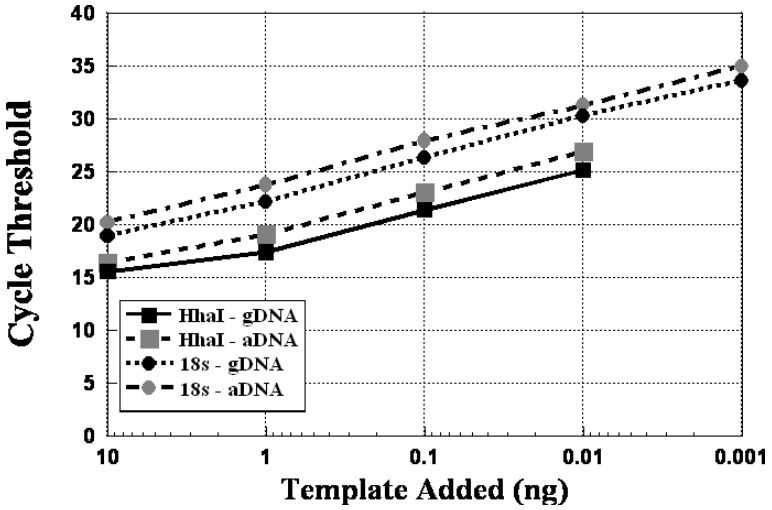


Figure 1: Results of q-PCR assays comparing genomic and amplified DNA. q-PCR assays targeted the HhaI tandem repeat or the 18s rDNA of genomic DNA (gDNA) and DNA amplified from 10 ng of gDNA template (aDNA).

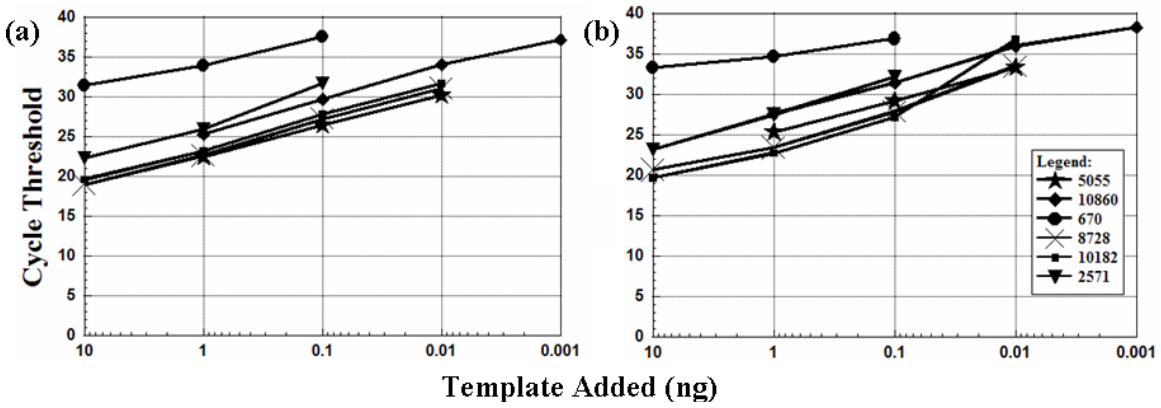


Figure 2: Results of SybrGreen q-PCR assays targeting single and low copy number *B. malayi* genes. Templates include: a cysteinyl tRNA synthetase (5055), three hypothetical proteins (10860, 670, 2571), a Rab family protein (8728) and a putative serine/threonine phosphatase (10182). (a) Represents gDNA while (b) represents aDNA.

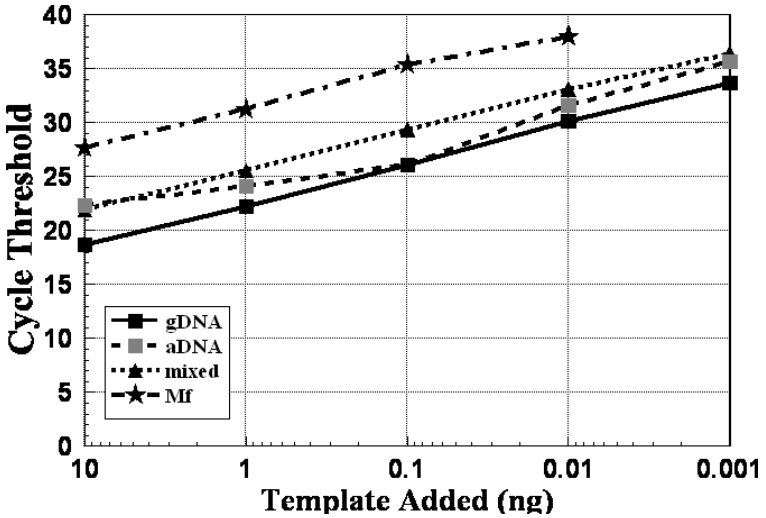


Figure 3: Results of q-PCR reactions with comparing various templates. Templates include unamplified gDNA (gDNA), aDNA made from 0.5 ng of gDNA template (aDNA), aDNA made from a template mixture of 5.0 ng of *B. malayi* gDNA and 5.0 ng human DNA (mixed) and aDNA made using 1/15 of the lysate of one microfilaria as template (MF).

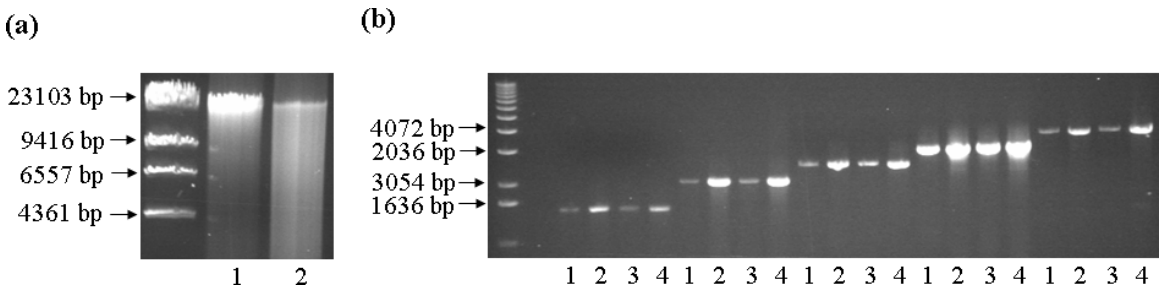


Figure 4: Assessment of the size of aDNA fragments by gel electrophoresis (a) and conventional PCR analysis (b). (a) Gel electrophoresis indicates that the largest DNA fragments found in aDNA (lane 2) are comparable in size to those found in gDNA (lane 1). One microgram of DNA was loaded in lanes 1 and 2. (b) Conventional PCR reactions were able to amplify products ranging in size from 1505 to 3585 bp from 1.0 ng and 10 ng of gDNA template (lanes 1 and 2, respectively) and from 1.0 at 10 ng of aDNA template (lanes 3 and 4, respectively). PCR primers used in these reactions are described in **Table 1**.

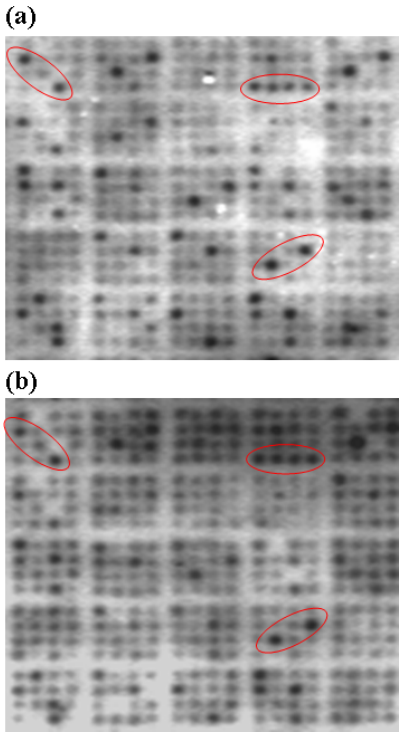


Figure 5: Identical representative sections of the BAC array hybridized with labeled gDNA (a) and aDNA (b). Hybridization patterns obtained with these two DNA types had > 90% agreement. Circles highlight examples of agreement.

TABLES

Table 1: Sequences of oligonucleotides used to detect fragments of the 18S rDNA of *B. malayi* by q-PCR (amplicon size 120 bp) or fragments of the X-chromosome (amplicon sizes 1,505 – 3,585) by conventional PCR.

Product Size (BP)	Forward Primer (5'-3')	Reverse Primer (5'-3')
120	CGAATTTTAAACGAGCGCATCT / *6Fam*ATTTAATCCGTTAATTGGTGACTC*MGBNFQ**	CATGCGATCAGCCATAGCTATT
1505	CGGGGTTTCAGGTATTTGTGG	ATTTGGCGGTAAGCTTGTTGTCT
1896	CTGCTGCTAATGACTGTCTG	GGAAGGTCTGAAACCGATTG
2507	AACGCTTGGCTTCCCTAAT	CAACTCCCCTGTCGTCA
3021	GGGTTCCTCTGATGCTACTTTC	CTTGATTCGACACCGATGAC
3585	GTTCGAGCAGCACTTGTCATAGG	GAAACCATCATCGGCTCGTCT

* TaqMan probe

Table 2: Primer sequences and targets of SybrGreen q-PCRs analyzing single and low copy number protein coding genes (based on genome data by Ghedin et al. [10]. The size of the q-PCR product was 65 bp.

EST Accession #	# scaffolds containing EST sequence:	Description	Forward Primer (5'-3')	Reverse Primer (5'-3')
AA072504	0	Cysteinyl tRNA synthetase	TGAGATAGTAATTTGCAGTCCATC	AGTGCGCTAACGTAAGGCATAAT
AW203926	0	Hypothetical protein	CGCGAAACATGTGTAAAAGCAT	CCATTTTCAAAAGGTGAATGACATT
AI943531	1	Hypothetical protein	CAACGATTTCGCGATCATGTT	GCATTTTGACGATTTCTCGCTAA
AI783139	1	Rab family protein	AATAAAGAACCGAATGGGTCCAA	CCCCCTAATGCGCACTGAT
AI784778	1	Putative serine/threonine phosphatase	GGAAGGTGGTGGATTGACTAAAAA	GAGCTTTCATACACAAAGCAAGCA
AA471542	1	Hypothetical protein	TCACCGTCGAAATCTGCTTCT	CGCGTGATGAGTTATCAGATGAA

Table 3: Hybridization of gDNA and aDNA to a filarial oligonucleotide microarray. Percentages shown are mean percentages of flags detected in hybridizations to probes with sequences from the indicated species observed with either gDNA or aDNA made from 10 ng gDNA template. *P* values for differences between native genomic and amplified *B. malayi* DNA (assessed by Student's T-test) were: 0.19 for *B. malayi* clusters, 0.21 for *O. volvulus* clusters, 0.37 for *W. bancrofti* clusters and 0.13 for *Wolbachia B.m.* clusters. N values indicate the number of samples tested.

Species	No. Clusters	% of flags detected		
		genomic *B. malayi* DNA (n=7)	Amplified *B. malayi* DNA (n=10)	No Template Control (n=1)
B. malayi	15412	94.99 ± 1.27	93.30 ± 3.52	11.37
O. volvulus	1016	50.82 ± 6.82	55.70 ± 8.62	11.12
W. bancrofti	872	72.30 ± 4.42	74.50 ± 5.30	16.40
Wolbachia B.m.	803	99.45 ± 1.17	92.81 ± 12.58	4.73